

ATP Synthases: Insights Into Their Motor Functions From Sequence and Structural Analyses

Sangjin Hong¹ and Peter L. Pedersen^{1,2}

Received January 3, 2003; accepted February 1, 2003

ATP synthases are motor complexes comprised of F_0 and F_1 parts that couple the proton gradient across the membrane to the synthesis of ATP by rotary catalysis. Although a great deal of information has been accumulated regarding the structure and function of ATP synthases, their motor functions are not fully understood. For this reason, we performed the alignments and analyses of the protein sequences comprising the core of the ATP synthase motor complex, and examined carefully the locations of the conserved residues in the subunit structures of ATP synthases. A summary of the findings from this bioinformatic study is as follows. First, we found that four conserved regions in the sequence of γ subunit are clustered into three patches in its structure. The interactions of these conserved patches with the α and β subunits are likely to be critical for energy coupling and catalytic activity of the ATP synthase. Second, we located a four-residue cluster at the N-terminal domain of mitochondrial OSCP or bacterial (or chloroplast) δ subunit which may be critical for the binding of these subunits to F_1 . Third, from the localizations of conserved residues in the subunits comprising the rotors of ATP synthases, we suggest that the conserved interaction site at the interface of subunit c and δ (mitochondria) or ϵ (bacteria and chloroplasts) may be important for connecting the rotor of F_1 to the rotor of F_0 . Finally, we found the sequence of mitochondrial subunit b to be highly conserved, significantly longer than bacterial subunit b , and to contain a shorter dimerization domain than that of the bacterial protein. It is suggested that the different properties of mitochondrial subunit b may be necessary for interaction with other proteins, e.g., the supernumerary subunits.

KEY WORDS: F_0F_1 ATP synthase; biological motor; protein sequence alignment; bioinformatics.

INTRODUCTION

ATP synthases are rotary motor complexes that play a central role in oxidative or photosynthetic phosphorylation, coupling the flow of protons down an electrochemical gradient to the synthesis of ATP. ATP synthases are composed of two discrete sectors (F_1 and F_0) that are considered to be separate rotary motors working cooperatively (for reviews, see Capaldi and Aggeler, 2002; Pedersen *et al.*, 2000a). The essential part of the F_1 motor is an ATP-driven $\alpha_3\beta_3\gamma$ subcomplex in which the rotor (γ subunit) is

held by a stator $\alpha\beta$ trimer. In the core unit of the F_0 motor that is composed of ac_{10-14} , the proton-driven subunit c ring rotor rotates relative to the stator (a subunit).

The simplest ATP synthases, i.e. those from nonphotosynthetic eubacterial sources are composed of eight different subunit types (Fillingame and Divall, 1999), which appear to be the minimal composition required for activity. The chloroplast and photosynthetic bacterial ATP synthases consist of nine different subunit types (Borghese *et al.*, 1998; Richter *et al.*, 2000). The mitochondrial ATP synthase is more complicated, as the animal and yeast (or fungal) enzymes are reported so far to consist of 16 and 20 different subunit types, respectively (Pedersen *et al.*, 2000b; Velours and Arselin, 2000). In addition, a potential regulator named STF₃ has been identified recently in yeast from a Blast search of the yeast genome and EST database (Hong and Pedersen, 2002) (Fig. 1).

¹ Department of Biological Chemistry, School of Medicine, Johns Hopkins University, 725 N. Wolfe Street, Baltimore, Maryland 21205-2185.

² To whom correspondence should be addressed; e-mail: ppedersen@jhmi.edu.

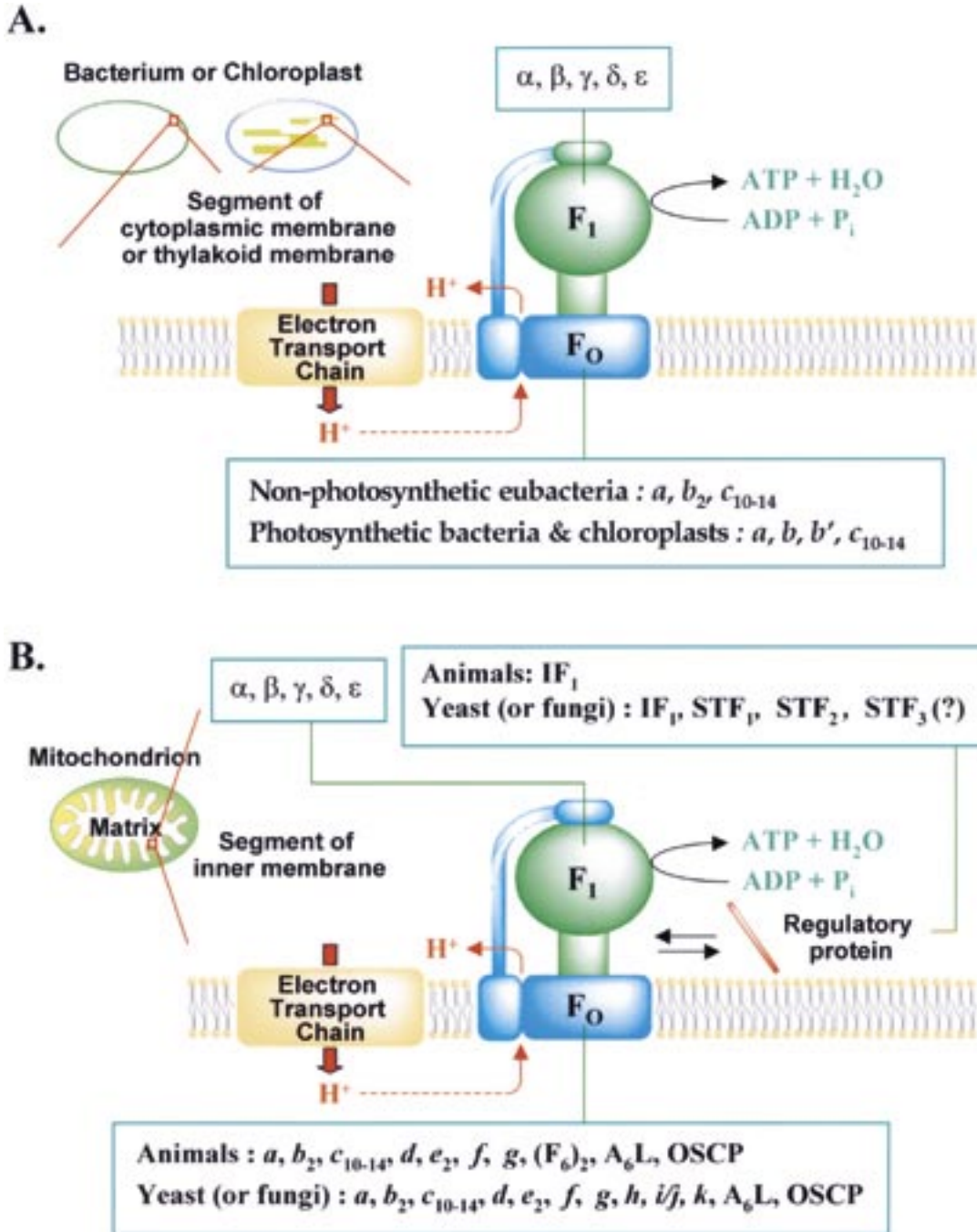


Fig. 1. Schematic structure of F₀F₁ ATP synthase from bacteria and chloroplasts (A) and mitochondria (B).

During the last decade, great strides have been made in understanding the structure of ATP synthases, a result in part from the convergence of biochemical and biophysical studies that have provided valuable information about the structures of these enzymes. Recently, the visualization

of the rotation of F₁ using a video microscopic technique provided direct evidence for the rotation of the γ subunit within the F₁ motor of an ATP synthase (Noji *et al.*, 1997; Tanabe *et al.*, 2001). Despite this and other remarkable progress, much more work needs to be carried out to

fully understand the complicated motor function of ATP synthase. To gain structural insight into the rotary motor functions of ATP synthases, we performed alignments and analyses of the protein sequences that comprise the motor components of ATP synthases. From these studies together with an examination of conserved residues and structural motifs within the available X-ray crystallographic and NMR structures, we found several intriguing structural features that may be related to the motor functions of ATP synthases.

MATERIALS AND METHODS

The protein sequences of the subunits of ATP synthase were retrieved using PSI-BLAST database search (Altschul *et al.*, 1997; Altschul and Konini, 1998) with the default parameters (BLOSUM62 matrix, 0.005 as an E-value threshold, no low complexity filtering in the query sequence) on the nonredundant (nr) protein database. The protein sequences were aligned by the programs ClustalW (Thompson, 1994) and Macaw (Schuler *et al.*, 1991) followed by manual adjustment. The sequence analyses for the prediction of secondary structure, transmembrane helices, and structural motifs were performed using pSAAM (Crofts, 1992), Peptool (Wishart *et al.*, 2000), OMIGA (Kramer, 2001), SOSUI (Hirokawa *et al.*, 1998), TMHMM (Krogh *et al.*, 2001), Coiled-coil prediction server (Lupas *et al.*, 1991), and PredicProtein server (<http://www.emblheidelberg.de/predictprotein/predictprotein.html>). The programs WebLab ViewerPro and Quanta were used for the representation of protein structures. The full names of the species of which the protein sequences are used in this study are [Fungi] *A. bisporus*—*Agaricus bisporus*, *K. lactis*—*Kluyveromyces lactis*, *N. crassa*—*Neurospora crassa*, *S. pombe*—*Schizosaccharomyces pombe*, Yeast—*Saccharomyces cerevisiae*; [Euglenozoa] *E. gracilis*—*Euglena gracilis*; [Metazoa] African clawed frog—*Xenopus laevis*, Bovine—*Bos taurus*, *C. elegans*—*Caenorhabditis elegans*, Fruit fly—*Drosophila melanogaster*, Human—*Homo sapiens*, Mouse—*Mus musculus*, Rat—*Rattus norvegicus*; [Haptophyceae] *O. neapolitana*—*Ochrosphaera neapolitana*; [Red Algae] *Antithamnion sp.*, *C. caldarium*—*Cyanidium caldarium*, *P. purpurea*—*Porphyra purpurea*; [Stramenopiles] *O. sinensis*—*Odontella sinensis*, *P. littoralis*—*Pylaiella littoralis*, *P. tricornerutum*—*Phaeodactylum tricornerutum*, [Cryptomonads] *G. theta*—*Guillardia theta*; [Green plants] *A. columnaris*—*Aegilops columnaris*, Alfalfa—*Medicago sativa*, *A. thaliana*—*Arabidopsis thaliana*, Barley—

Hordeum vulgare, *C. reinhardtii*—*Chlamydomonas reinhardtii*, *C. vulgaris*—*Chlorella vulgaris*, Japanese black pine—*Pinus thunbergii*, Liverwort—*Marchantia polymorpha*, Maize—*Zea mays*, Norway spruce—*Picea abies*, Pea—*Pisum sativum*, Pineapple—*Ananas comosus*, Rice—*Oriza sativa*, Sorghum—*Sorghum bicolor*, Spinach—*Spinacia oleracea*, Sweet potato—*Ipomoea batatas*, Tobacco—*Nicotiana tabacum*, Wheat—*Triticum aestivum*; [Bacterial] *A. aeolicus*—*Aquifex aeolicus*, *A. ferrooxidans*—*Acidithiobacillus ferrooxidans*, *Anabaena PCC7120*, *A. tumefaciens*—*Agrobacterium tumefaciens*, *Bacillus sp. PS3*, *B. aphidicola*—*Buchnera aphidicola*, *B. caldotenax*—*Bacillus caldotenax*, *B. firmus*—*Bacillus firmus*, *B. halodurans*—*Bacillus halodurans*, *B. megaterium*—*Bacillus megaterium*, *B. pseudofirmus*—*Bacillus pseudofirmus*, *B. stearothermophilus*—*Bacillus stearothermophilus*, *B. subtilis*—*Bacillus subtilis*, *Buchnera sp. APS*, *C. acetobutylicum*—*Clostridium acetobutylicum*, *C. jejuni*—*Campylobacter jejuni*, *E. coli*—*Escherichia coli*, *E. hirae*—*Enterococcus hirae*, *H. influenzae*—*Haemophilus influenzae*, *H. pylori*—*Helicobacter pylori*, *L. acidophilus*—*Lactobacillus acidophilus*, *M. gallisepticum*—*Mycoplasma gallisepticum*, *M. genitalium*—*Mycoplasma genitalium*, *M. leprae*—*Mycobacterium leprae*, *M. thermoacetica*—*Moorella thermoacetica*, *M. tuberculosis*—*Mycobacterium tuberculosis*, *N. meningitidis Z2491*—*Neisseria meningitidis Z2491*, *P. aeruginosa*—*Pseudomonas aeruginosa*, *P. didemni*—*Prochloron didemni*, *P. modestum*—*Propionigenium modestum*, *P. multocida*—*Pasteurella multocida*, *R. blasticus*—*Rhodobacter blasticus*, *R. capsulatus*—*Rhodobacter capsulatus*, *R. prowazekii*—*Rickettsia prowazekii*, *R. rubrum*—*Rhodospirillum rubrum*, *S. bovis*—*Streptococcus bovis*, *S. lividans*—*Streptomyces lividans*, *S. mutans*—*Streptococcus mutans*, *S. typhimurium*—*Salmonella typhimurium*, *Synechococcus PCC6716*, *Synechocystis PCC6803*, *T. ferrooxidans*—*Thiobacillus ferrooxidans*, *T. maritima*—*Thermotoga maritima*, *V. alginolyticus*—*Vibrio alginolyticus*, *V. cholerae*—*Vibro cholerae*, *X. fastidiosa*—*Xylella fastidiosa*.

RESULTS AND DISCUSSION

Interaction of the γ Subunit With the α and β Subunits

The sequence alignment of the γ subunit shows that the overall sequence is not that conserved in comparison to the α and β subunits although subunit γ is one of the

essential subunits for the function of the F_1 motor (Fig. 2). In the sequence alignments of the γ subunits from different sources, the completely and highly conserved residues are distributed fairly clustered, and the clustered regions are primarily localized in the upper and middle regions of the paddle-like structure of this subunit. The lower part of the γ structure as it protrudes from the bottom of F_1 comprises the middle part of the sequence. It contains few completely or highly conserved residues. The residues conserved completely or completely with single variation are found in more than 95% of the 27 protein sequences analyzed. These are clustered into four regions in the sequence of γ (I, 19–26; II, 85–88; III, 238–253; IV, 264–282 in *E. coli* sequence). It was previously suggested from a suppression mutagenesis study that three segments in the sequence of γ subunit (18–35, 236–246, 269–280 in *E. coli* sequence) were critically involved in the energy coupling and catalytic activity of ATP synthase (Nakamoto and Al-Shawi, 1995). The three segments suggested in the paper correspond to I, III, IV regions in this study. The region II, which is unique in this study, is also believed to be important for the function of ATP synthases. The four conserved regions of the γ subunit are rearranged into three patches in the structure of γ . Three of the conserved regions, I, II, and III, congregate into a conserved patch (A) in the middle of the γ structure, and region IV segregates into two patches (B and C) at the upper part of γ (Fig. 3(A)).

In α and β subunits, on the other hand, the residues conserved completely or completely with single variation are largely localized on the side facing the γ subunit, covered by nonconserved and weakly conserved residues. Among the conserved α and β residues, those located at the interface with the γ subunit directly interact with the three conserved patches of this subunit (Fig. 3(B)). The conserved regions of α and β subunits from residues S274 to G288, and E253 to L271 in the *E. coli* sequence, respectively, interact with the conserved patch C of the γ subunit. The conserved residues G335 to F340 of α and Y297 to T310 of β interact with the conserved patch B of the γ subunit. Residues A405 to D414 of α and D372 to K387 of β located in the α -helical domains of these subunits interact with the conserved patch A of γ . Significantly, the three conserved patches of the γ subunit interact differently with the three different conformational forms of the α and β subunits (TP, DP, and E) (Fig. 3(B)). The localized distribution of conserved residues in the γ subunit and the differential patterns of interactions of the three conserved patches of this subunit with the α and β subunits strongly suggest that these specific interactions may be crucial for the rotary motor function of F_1 .

The Potential Interaction Site of OSCP (Mitochondria) or δ (Bacteria and Chloroplasts) With F_1

It has been reported that the mutation of the residue Arg94 of rat OSCP caused defects in the interactions of OSCP with F_1 (Golden and Pedersen, 1998). Another study with bovine OSCP showed the same mutational effects when the N-terminal 28 amino acid residues of bovine OSCP were deleted (Joshi *et al.*, 1996). In an attempt to better define the potential interaction sites of OSCP with F_1 , we aligned the sequences of OSCP with those of the δ subunit (OSCP equivalent) from bacteria and chloroplasts, and then examined carefully the conserved residues in the structure of the δ subunit's N-terminal domain (1abv; Wilkens *et al.*, 1997).

In the N-terminal domain structure of the bacterial δ subunit, the region corresponding to the N-terminal 28 residues in OSCP, forms an N-terminal α -helix, and residue 84, which corresponds to the residue 94 in OSCP, is located within a loop connecting two other helices (Fig. 4(A)). In the sequence alignment of OSCP (mitochondria) and the δ subunit (bacteria and chloroplasts), the N-terminal domain structure of the bacterial δ subunit includes four highly conserved residues. Three of these (Tyr10, Ala11, Ala13 in the *E. coli* sequence) are seen in the N-terminal α -helix (Fig. 5). The other conserved residue in the N-terminal domain structure, Arg84 in the *E. coli* sequence, is positioned closely to the three conserved residues in the N-terminal α -helix. The detailed examination of the four residues in the N-terminal domain structure lead us to find that the residues Tyr10, Ala13, and Arg84 are aligned in a row on the surface of the structure (Fig. 4(B)). The residue Ala11 is located within the molecule behind the three residues. This four-residue cluster is the only conserved site in the 105 amino acid residue N-terminal domain. Most of the weakly conserved residues within the domain are also positioned around the four-residue conserved cluster. From the arrangement of conserved residues, we suggest that the four-residue conserved cluster plays an essential role in the interaction of OSCP or the bacterial/chloroplast δ subunit with F_1 (Fig. 4(C)).

Specific Interaction Sites at the Interface of the F_1 and F_0 Rotors

In ATP synthases, the rotor of the F_1 motor and the rotor of the F_0 motor are generally considered to be comprised respectively of the $\gamma\epsilon$ subunit pair (or $\gamma\delta$ in mitochondria) and the c subunit ring. The two rotors have been

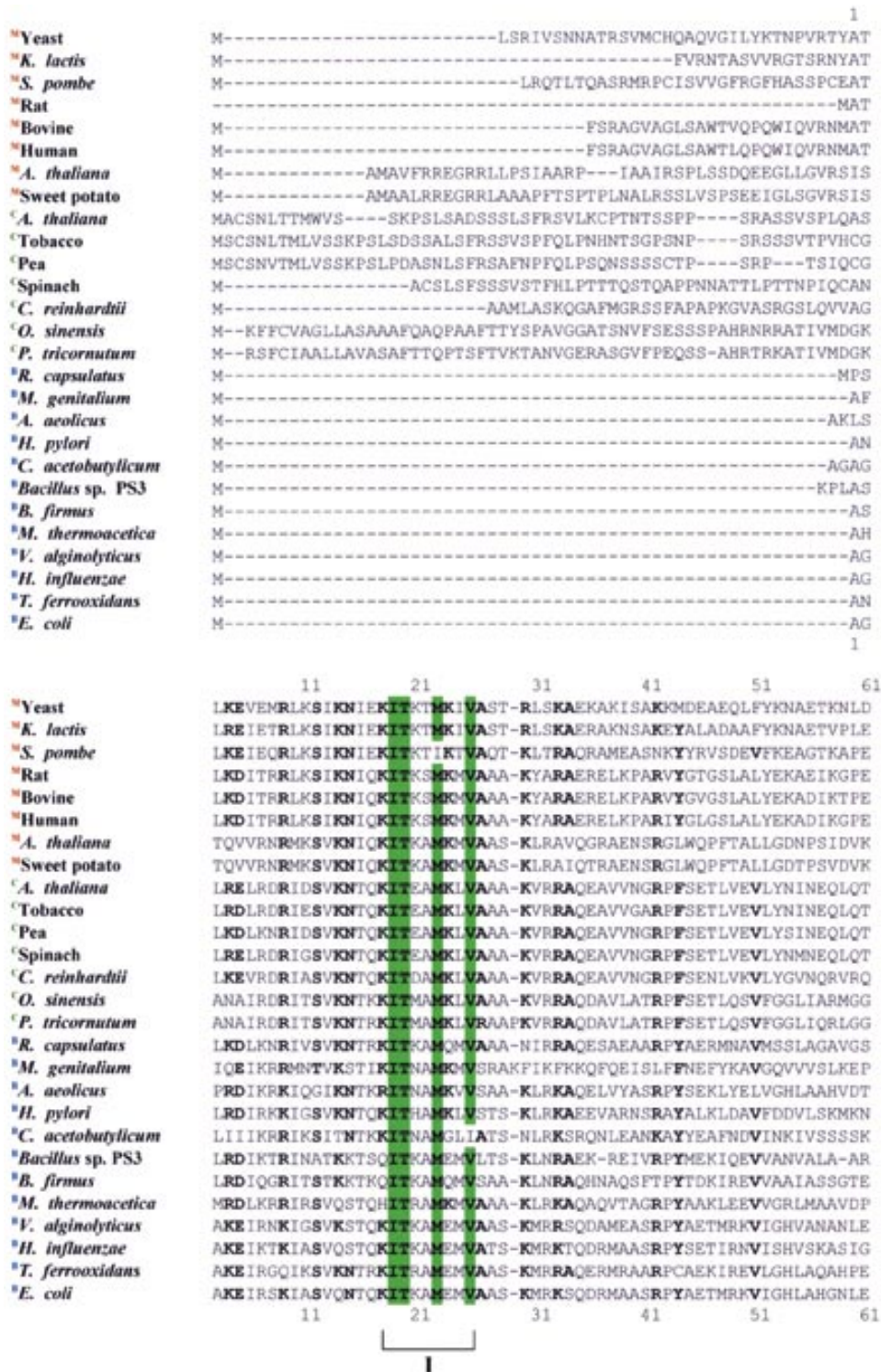


Fig. 2. Sequence alignment of the γ subunit. The conserved residues are in bold, and the residues conserved completely, and completely with only a single deviation, are highlighted in shade. Note that the residues conserved completely or completely with a single variation are clustered in four regions in the sequence of γ (I, II, III, and IV). The letters M, C, and B in superscript denote the source of the sequence; mitochondria, chloroplast, and bacteria, respectively.

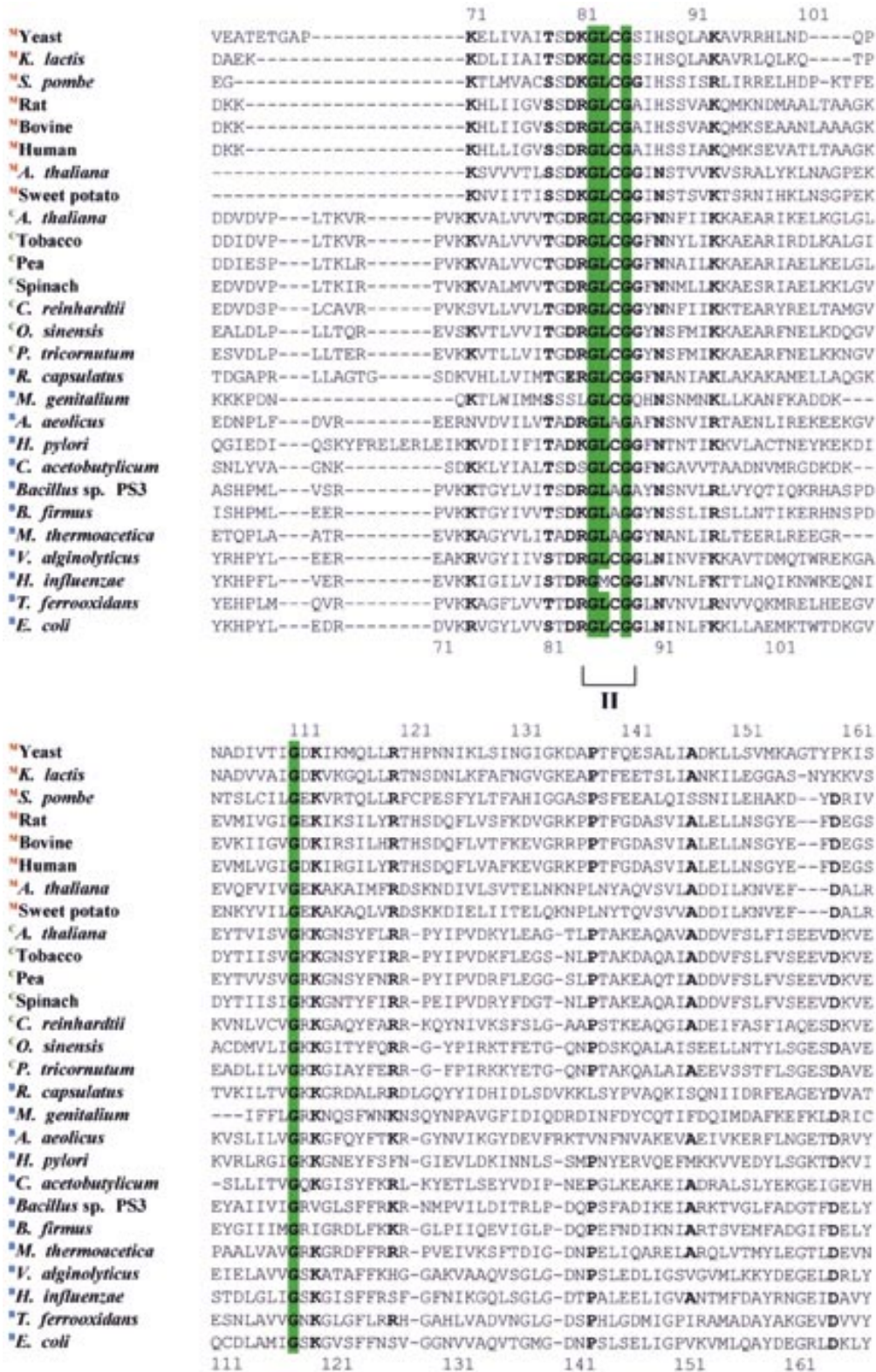


Fig. 2. Continued

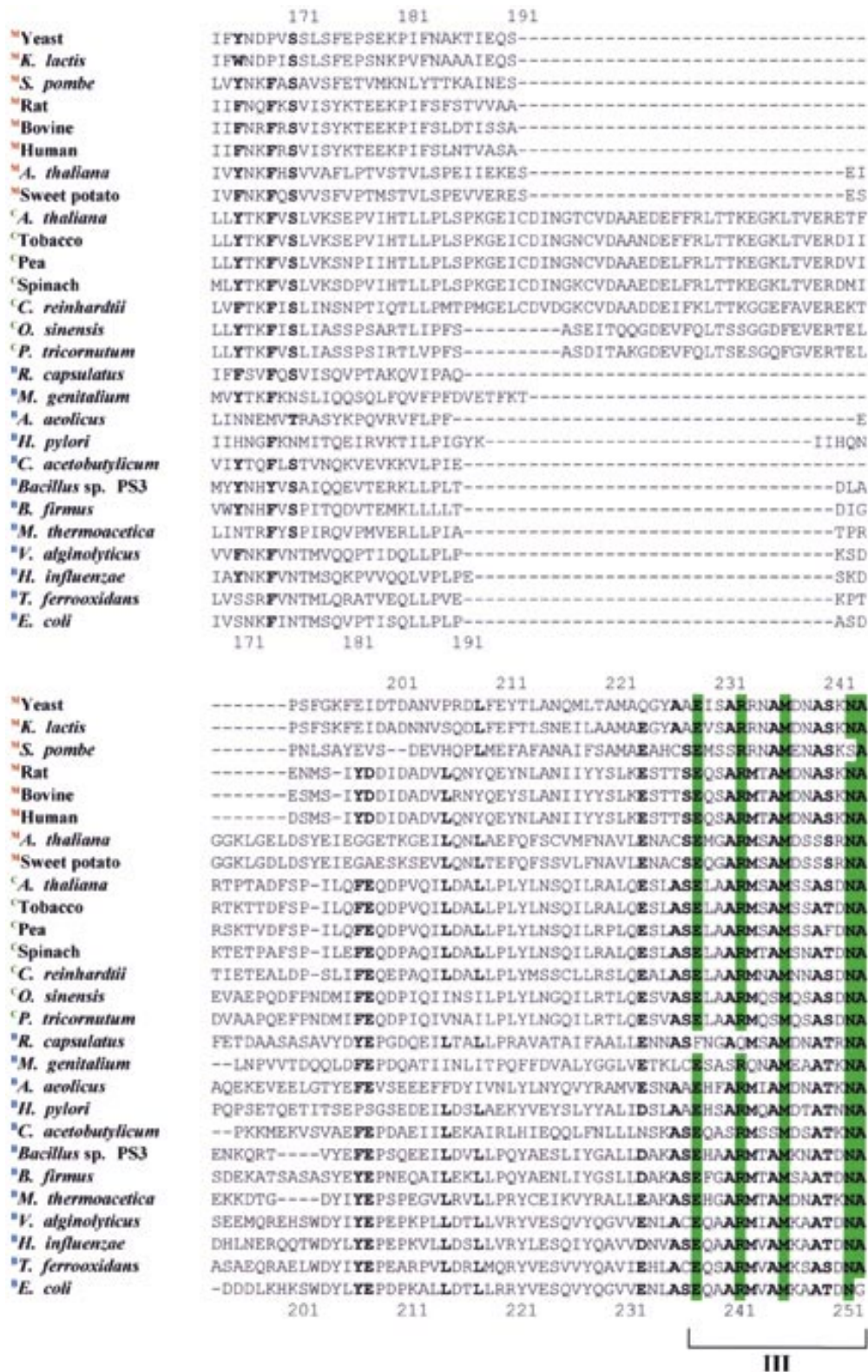


Fig. 2. Continued

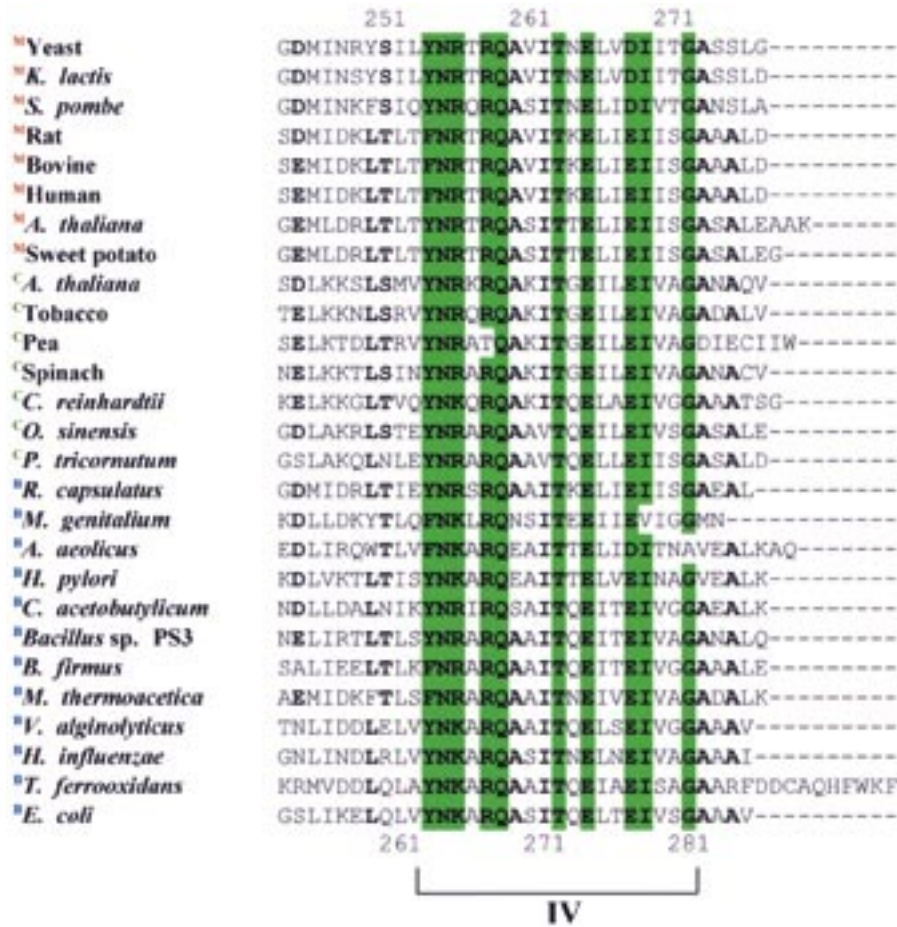


Fig. 2. Continued

reported to behave as one mechanical unit in ATP synthase (Tanabe *et al.*, 2001). In mitochondrial ATP synthase, the ϵ subunit, the counterpart of which is not present in bacterial or chloroplast ATP synthase, may also be a part of the F_1 rotor. Experimentally, the two motors of ATP synthase, F_1 and F_0 can be easily separated, and also can be reconstituted to form a functional F_0F_1 motor complex. The reversibility of the strip of F_1 , and reconstitution of F_1 and F_0 , indicate that the F_1 rotor and the F_0 rotor have specific sites for interaction. On the basis of the current available X-ray crystal structural data of ATP synthase, one of the potential specific interaction sites is believed to be located at the interfaces between the rotors of F_1 and F_0 where the $\gamma\epsilon$ (bacteria and chloroplasts) or $\gamma\delta$ (mitochondria) subunits of F_1 rotor have direct contacts with the subunit c ring comprising the F_0 rotor. To gain insight into the potential specific interaction site at the interface

of the rotors of F_1 and F_0 , we aligned the sequences of the three subunit types, γ , δ (mitochondria) or ϵ (bacteria and chloroplasts), and c subunits, and examined the conserved residues at the interfaces of the subunits in the available crystal structures.

The sequence alignment of subunit c shows that seven residues are completely or highly conserved, and three of them (Arg41, Gln42 or Asn, and Pro43 in *E. coli* sequence) are located facing F_1 at the interhelical loop (Fig. 6 and 8(A)). In the δ (mitochondria) or ϵ (bacteria and chloroplasts) subunit where the sequence alignments show five completely or highly conserved residues all positioned in the N-terminal β -sandwich domain (residues from 1 to 84 in the *E. coli* sequence) (Fig. 7 and Fig. 8(B)), three completely or highly conserved residues (G43, G46, and H51 in the bovine sequence) are localized at the interface with subunit c facing the interhelical loop of this subunit

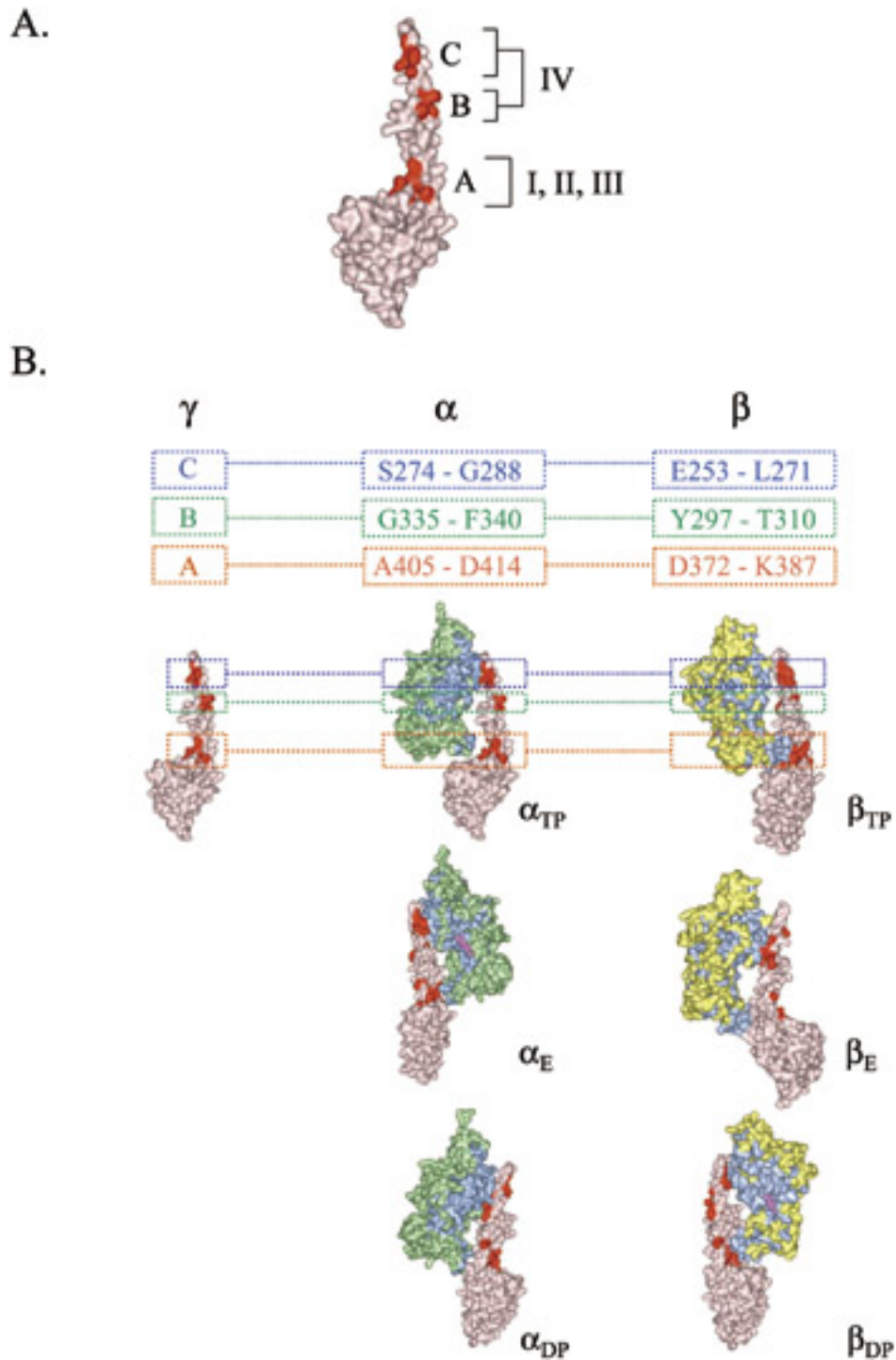


Fig. 3. Three conserved patches in the structure of the γ subunit (1e79; Gibbons *et al.*, 2000). (A) Four conserved regions in the sequence of the γ subunit are clustered into three patches (red) in the structure. (B) Each conserved patch of the γ subunit interacts with the conserved regions of the α and β subunits in different patterns depending on the conformations of α and β subunits.

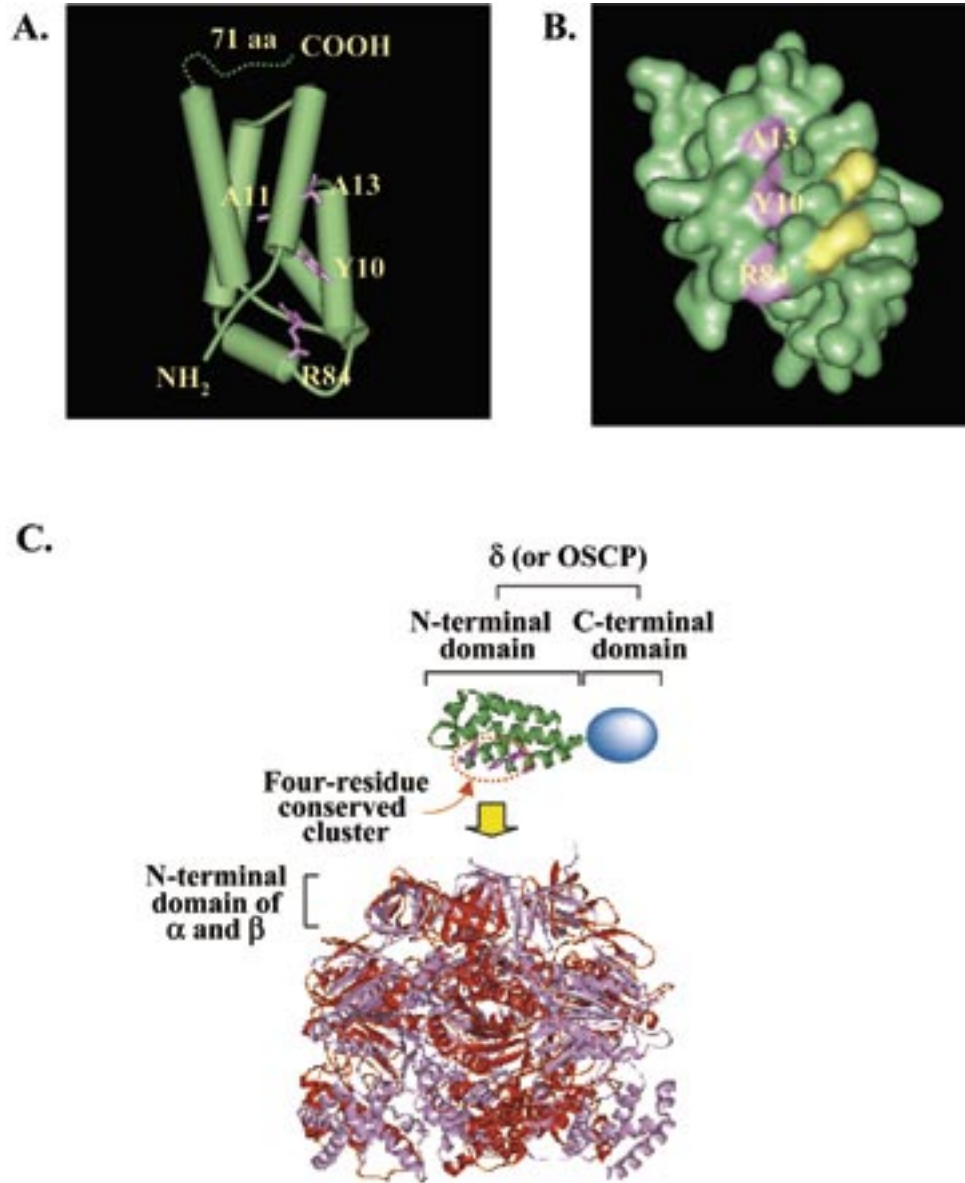


Fig. 4. Location of the highly conserved four-residue cluster in the structure of the N-terminal domain of the bacterial δ subunit, shown in cylinder (A) and surface (B) representations. The residues that form a four-residue conserved cluster in the N-terminal domain, drawn in purple, are located on one side of the molecule, and three of the residues are arrayed in a line on the surface of the molecule. Weakly conserved residues, colored in yellow, are also localized on one side of the molecule around the four-residue conserved cluster. (C) Potential orientation of the conserved cluster of the N-terminal domain of the bacterial δ subunit or mitochondrial OSCP with respect to the $\alpha\beta$ trimer.

(Fig. 8(B) and (C)). On the other hand, the sequence alignment of the γ subunits, and a search for the conserved residues in the structure, show no completely or highly conserved residues at the interface of γ and c subunits. Therefore, taken together, we suggest that interactions between subunit c and the δ subunit are directly involved in

the reconstitution process of F_1 and F_0 , and that the conserved residues at the interface of the two subunits [Arg41, Gln42 or Asn, and Pro43 in subunit c (*E. coli* sequence) and G43, G46, and H51 in mitochondrial δ or the bacterial ε subunit (bovine sequence)], play a key role in rotor connections.

Mitochondrial OSCP

³⁴ Yeast	M-----FNRVFTRSF-----
³⁴ <i>S. pombe</i>	M-----NHIFRRSIPITARLPS-----
³⁴ <i>K. lactis</i>	M-----FSRVFVR-----
³⁴ Fruit fly	M-----ASINKLALL-----
³⁴ Rat	M-----AAPATSVLSRQVRS-----
³⁴ Bovine	M-----AALAVSGLSQVRC-----
³⁴ Human	M-----AAPAVSGLSRQVRC-----
³⁴ <i>A. thaliana</i>	M-----ANRFRSGISFFKTIAVTDSVSSVRSKSLFP-----A
³⁴ Sweet potato	M-----AMTGRARSMGFSILQKALSSAQRSNAHRSLCPTLSNSEL

Bacterial and Chloroplast δ

^C <i>P. purpurea</i>	M-----
^C <i>G. theta</i>	M-----
^C <i>O. neapolitana</i>	M-----
^C <i>Antithamnion</i> sp.	M-----
^C <i>O. sinensis</i>	M-----
^C <i>C. caldarium</i>	M-----
^C Tobacco	M----MAALQQTPIITFQSRSPPTQIIISGPTAKLSFSGGLKLP---KLT
^C Pea	M----MASLQHTTASLHSHKHIPKTTNILTRKPIILNLSSSTFYSPKLLK
^C Spinach	MAALQNPVALQSRRTTAVAAALSTSTTSTPKPFSLSFSSSTATFNPLRLK
^C Sorghum	M----AALRLASFTLRPAAAAAASASSGATFAAPRSASFARAARGLPS
^C <i>C. reinhardtii</i>	M-----LAAKSIAGPRAFKASAVRAAPKAGRRTVVVMAR-----
^B <i>Anabaena</i> sp. PCC 7120	M-----
^B <i>Synechocystis</i> sp. PCC 6803	M-----
^B <i>Synechococcus</i> sp. PCC 6301	M-----
^B <i>M. genitalium</i>	M-----
^B <i>M. gallisepticum</i>	M-----
^B <i>M. thermoacetica</i>	M-----
^B <i>P. modestum</i> *	M-----
^B <i>C. acetobutylicum</i>	M-----
^B <i>S. bovis</i>	M-----
^B <i>Buchnera</i> sp. APS	M-----
^B <i>R. prowazekii</i>	M-----
^B <i>T. maritima</i>	M-----
^B <i>C. jejuni</i>	M-----
^B <i>H. pylori</i>	M-----
^B <i>S. typhimurium</i>	M-----
^B <i>V. cholerae</i>	M-----
^B <i>V. alginolyticus</i>	M-----
^B <i>H. influenzae</i>	M-----
^B <i>P. multocida</i>	M-----
^B <i>P. aeruginosa</i>	M-----
^B <i>X. fastidiosa</i>	M-----
^B <i>T. ferrooxidans</i>	M-----
^B <i>A. ferrooxidans</i>	M-----
^B <i>N. meningitidis</i> Z2491	M-----
^B <i>R. capsulatus</i>	M-----
^B <i>R. blasticus</i>	M-----
^B <i>A. tumefaciens</i>	M-----
^B <i>Bacillus</i> sp. PS3	M-----
^B <i>B. subtilis</i>	M-----
^B <i>B. firmus</i>	M-----
^B <i>B. pseudofirmus</i>	M-----
^B <i>B. halodurans</i>	M-----
^B <i>L. acidophilus</i>	M-----
^B <i>E. coli</i>	M-----

Fig. 5. Sequence alignment of OSCP (mitochondria) and δ (bacteria and chloroplasts) subunits. Highly and weakly conserved residues are highlighted in bold and shaded in purple and light yellow, respectively. The residues that form a four-residue conserved cluster in the N-terminal domain (residues 1 to 105 in the *E. coli* sequence) are indicated by arrows. * *P. modestum* is Na⁺-ion specific.

Mitochondrial OSCP

	1	11	21	31
³⁴ Yeast	--ASSLRAAASKAAAPPVRLFGVEGT	FATL	LYQAAA	--KNSSIDAAPQSL
³⁴ <i>S. pombe</i>	--LGIRSLATATASAHPPVQLYGLDGS	FASS	LYTAAV	--KESKLDNVEKAL
³⁴ <i>K. lactis</i>	-----SMAASAKAGVKPPVQLFGLDGT	FATL	LPTAAS	--KTTSVAAAASSL
³⁴ Fruit fly	--SRTLSSAAAQATVKPPVQVFLGEGRE	FATAL	LYSAAS	--KLSQLDQVEKDL
³⁴ Rat	--FSTSVVRPFKLVLRPPVQVYGI	EGRE	FATAL	LYSAAS--KQKRLDQVEKEL
³⁴ Bovine	--FSTSVVRPFKLVLRPPVQIYGI	EGRE	FATAL	LYSAAS--KQNKLEQVEKEL
³⁴ Human	--FSTSVVRPFKLVLRPPVQVYGI	EGRE	FATAL	LYSAAS--KQNKLEQVEKEL
³⁴ <i>A. thaliana</i>	LRTYATASAQTTANVKVPIALVGENGN	PASW	LYIAAV	--KMNSLEKIEITDL
³⁴ Sweet potato	LRNYATASASKEQKIKVPLTMYGVSGN	PAS	LYLAAV	--KSNTLEKVESEL

Bacterial and Chloroplast δ

⁶ <i>P. purpurea</i>	-----SSNNLVAKIAQPTASALDLIAN	----	EKKAIEQISQDM
⁶ <i>G. theta</i>	-----IAMNKNLAQPTAMAFLEFSLDAK	QTLDTT	IADL
⁶ <i>O. neapolitana</i>	-----SVAKIADAEAEALLELIAN	----	SNKSLKETTNDM
⁶ <i>Antithamnion</i> sp.	-----SQSILYKIANPDAALLELSL	----	LNNAMDKASSDL
⁶ <i>O. sinensis</i>	-----SINPLASKIAAPFARALP	DFFSV	--DQNLMHQITADF
⁶ <i>C. caldarium</i>	-----HNNYKIAKVEEGLFEIAN	----	TKNSLAEINEQL
⁶ Tobacco	IKLRSNRTSRRGGGAAGSKMVASAAGS	TANLADI	AK--SNGTLEQTTADL
⁶ Pea	LKLPKTRRSTGGALGARMSSLAAGS	TAAALADI	AN--SNNTLDAITADF
⁶ Spinach	ILTASKLTAKPRGGALGTRMVDSTASR	FASALAD	VAD--VTGTLEATNSDV
⁶ Sorghum	LRLAPPFR-----GDLVRPRAEAAA	ADSASAL	SEVAV--ENGTLEQTVSDL
⁶ <i>C. reinhardtii</i>	-----KNEVSESAKALVELAD	----	EKGKLEAVHADV
¹⁰ <i>Anabaena</i> sp. PCC7120	-----TSKVANTEVAQPTAQALLSIAK	----	SKSLTEEPGTDA
¹⁰ <i>Synechocystis</i> sp. PCC 6803	-----KGSLSYSSKIAEPTAALIGLAQ	----	QQNLTEVFGONL
¹⁰ <i>Synechococcus</i> sp. PCC 6301	-----TSTSQFLDPTAEALMAIAR	----	EQGLEDRPGEA
¹⁰ <i>M. genitalium</i>	-----INAQAFGTALPQLSE	----	EQKQVKKIYECC
¹⁰ <i>M. gallisepticum</i>	-----DTNIMGFARALVDLAH	----	EEDKVHLYFYDNL
¹⁰ <i>M. thermoacetica</i>	-----SEQNVARRFARALFNIAH	----	EQGTAGEFANGL
¹⁰ <i>P. modestum</i>	-----IEAQVGRRAEAEIYEIAE	----	SNDNVKELYETL
¹⁰ <i>C. acetobutylicum</i>	-----YEFLLDRRALLYEVGE	----	KNQKLEEYINDF
¹⁰ <i>S. bovis</i>	-----DKKTQALVEQAKSLVEIAI	----	EKDSLAE-LQSE
¹⁰ <i>Buchnera</i> sp. APS	-----SVADTIARPTAQAIFI	FEIAI	--ENNTIEKW-KNI
¹⁰ <i>R. prowazekii</i>	-----NKDNLIGNFAVALFNNAI	----	LDNIQVKICEEI
¹⁰ <i>T. maritima</i>	-----RFSAVAGRFARALLNVAI	----	-----EKEKEEYL
¹⁰ <i>C. jejuni</i>	-----ENI IARRFAKAIASRAD	----	-----INDFYQN
¹⁰ <i>H. pylori</i>	-----QDLKVISKHEAKALKNHTK	----	SDLALLEEIVVG
¹⁰ <i>S. typhimurium</i>	-----SEFITVARPTAKA	AFDFAV	--EHQSVERW-QDM
¹⁰ <i>V. cholerae</i>	-----SDLTTIARPTAKA	AFDFAV	--EKQQLGHW-SQM
¹⁰ <i>V. alginolyticus</i>	-----SDLTTIARPTAKA	AFDFAL	--EKDQLDQW-GQM
¹⁰ <i>H. influenzae</i>	-----SELTTIARPTAKA	AFDFAI	--EQSAVEKW-TEM
¹⁰ <i>P. multocida</i>	-----SELTTIARPTAKA	AFDFAI	--EQNAVEKW-ANM
¹⁰ <i>P. aeruginosa</i>	-----AELTTLARPTAKA	AFEYAJ	--AHQQLADW-SAA
¹⁰ <i>X. fastidiosa</i>	-----SQALTLARPTAKA	AFAIAC	--EKGKCMQW-SQA
¹⁰ <i>T. ferrooxidans</i>	-----ADLITVARPTAEALMGWRK	----	--RAARNRPGRMH
¹⁰ <i>A. ferrooxidans</i>	-----ADLITVARPTAEALMGWRK	----	--RAARNRPGRMH
¹⁰ <i>N. meningitidis</i> Z2491	-----AEFATIAHPTAKALFGLAQ	----	EKSQIESW-LGG
¹⁰ <i>R. capsulatus</i>	-----SEPASISAAIAGREATAIFDLAQ	----	EAKGIDALSADV
¹⁰ <i>R. blasticus</i>	-----AEAASISQGI AERTATL	FELSK	--ETGALKTLETDI
¹⁰ <i>A. tumefaciens</i>	-----PVAETSQGTSGVAERTASSL	FELAL	--EAGTVEAVQVEL
¹⁰ <i>Bacillus</i> sp. PS3	-----NQEVIAKRTASL	FQIAL	--EQQLDRIEEDV
¹⁰ <i>B. subtilis</i>	-----SGSAVSKRTASL	FDIAN	--ESAQLNQVEEEL
¹⁰ <i>B. firmus</i>	-----SNQAVANRYRYALFQLAE	----	EKSILSQVVQEM
¹⁰ <i>B. pseudofirmus</i>	-----SNQAVANRYRYALFQLAE	----	EKSILSQVVQEM
¹⁰ <i>B. halodurans</i>	-----SNHAVANRYRYALFQLAE	----	EKGLHESPVSEL
¹⁰ <i>L. acidophilus</i>	-----ALSREEVAARNGTAL	PGYAQ	--DNKVLDTVYDEM
¹⁰ <i>E. coli</i>	-----SEFITVARPTAKA	AFDFAV	--EHQSVERW-QDM

Fig. 5. Continued

Mitochondrial OSCP

^MYeast QKVESTVKKPKLGHLLLN^PALS^LLKDRNSVIDAIVETHKNLDGYV^NLLK
^M*S. pombe* NKLSGVLQQRPEFEQYISS^PWLTREDKKILVSSLTQMTG-NEPLLNKFN
^M*K. lactis* NALKETVVKDTKLTSTILEN^PALS^AEDRNIVVDTLSEKIN-LDQSVQNLK
^MFruit fly TALQATIRSDKKLREYVTS^PIINKKVMATALKEASEKLR-FAPATVNLG
^MRat LRVGQLLK-DPKVSLAVLN^PYIKRSIKVKS^LKDITTEK-FSPLTANLMN
^MBovine LRVGQILK-EPKMAASLLN^PYVKS^VSVK^VKS^LSDMTAKEK-FSPLTANLN
^MHuman LRVAQILK-EPKVAASVLN^PYVKS^IIK^VKS^LNDITAKER-FSPLTANLN
^M*A. thaliana* SEMIEAMKTAPIFAQFTK^DPSVPRGTR^LAAIRDACDQAK-FAEPTKNFLS
^MSweet potato YDLVEASKKSP^TFSQFMRD^PSV^PVDTR^VNAI^KEICAQAK-FGDTTC^NFL

Bacterial and Chloroplast δ

^C*P. purpurea* KLIKIDILQSGKLYFLAN^PLK^TIEAKKQVIAATFGD-Q-ISENTLSFLM
^C*G. theta* TQIKTILHDSVDLSKTLN^PLLS^IKAKKEVIKAI^FEP-N-ISKNTLKFL
^C*O. neapolitana* NIVSQFLANSDDLKFLGN^PLITREK^KKNVLK^DVLGE-Q-ISSVSLN^FLM
^C*Antithamnion* sp. SMILEV^ISKSADLKL^FLSN^PLVEDN^LK^KNVLN^QL^FKD-K-VSD^FIVK^FLM
^C*O. sinensis* QNLEVF^LNKTPDLTEYLSN^PLISAKS^KEEVLN^KTLKS-Q-INKET^FKFLI
^C*C. caldarium* NSIKSVLQ^MPEFYFLVN^PLISQ^IK^KNTIKIV^FFN-N-LDKIT^LN^FLL
^CTobacco EKIEKISDD-EAVFNFFVS^PIVGEE^KKREL^VDEIVSSSS-IQPHVAN^FLN
^CPea DKIEQLFSD-PKVFDYFSS^PIVEDST^KRQLIGE^FATTSG-FQPHT^HN^FLN
^CSpinach EKLIRIFSE-EPVYFFAN^PVISIDN^KRSVL^{DEI}ITTS^G-LQPHTAN^FIN
^CSorghum EKLQKIFAD-ETVAEFFDN^PTVPREE^KTALIDEIA^KSYE-LQPHV^VN^FIN
^C*C. reinhardtii* DAVAGLMKENAKLSALIM^NPV^{ES}DKKRAVLAKIAKEAG-FQY^TIN^WLN
^B*Anabaena* sp. PCC 7120 RTLLNLLTENQQLRN^FIND^PPIAAEN^KKALIK^QILSE-A-SP-YLRN^FLL
^B*Synechocystis* sp. PCC 6803 RSLTLLQDSDLSAVLSS^PVVKDE^DK^KSVL^RSVLGD-G-GNGYLLN^FLM
^B*Synechococcus* sp. PCC 6301 ALFRSTLAASADLRHLEN^PTLFSS^QKAVLN^QVFGS-S-VHPLV^LN^FLN
^B*M. genitalium* HFFLKL^MMRNFKD^GLS^FLLNSY^TLTK^PDKIRL^VDKL^FKNH^FQ^VFV^DFLK
^B*M. gallisepticum* KVVFDLVKENQD--LMSLMNSQVLSKNQKHEI^DVV^FKDHL^TQ^TIV^DFLK
^B*M. thermoautotrophicum* BEVSR^TLAENSDFRRVLYHQ^LIPVRE^KQ^KLID^TIP^FPDIN---PLKN^FLN
^B*P. modestum* NGVMELYNTDKEFKTLVDH^PL^IKRED^KKEFAK^IFGELE---ESSINI^F
^B*C. acetobutylicum* GEIVHLLKNDENINQV^NH^PQISTSE^KK^IFMEIF^KGKI--DEKLLN^FLL
^B*S. bovis* TEALLSVFEETNLADFLSS^LVVSRDEK^VKL^VRLLQESS---V^MMN^FLE
^B*Buchnera* sp. APS LIFIKT^IASHK^KFK^NFLSG---SIS^PKYLS^LIFITIG^TNIIDENAK^NL^IK
^B*R. prowazekii* TLNSIIEDSFEIKK^FLS^PLVN^KIDKIN^VFN^SLVK^TTN-FNKIV^NN^FLL
^B*T. maritima* RFLDLVCQIYESSREL^FDN^PLKPE^KIS^LIKEIM^KSFGQEM^DEQ^FER^L
^B*C. jejuni* LCILNSAFVL^PK^FKN^IIESNE^IK^KER^KME^FLD^SFFD^IKN---SS^FQ^NFL^R
^B*H. pylori* LKNATEAIRQHKL^NQVLAHVSLK^VKEV^VFEILEKITS^IKAC^SV^LK^PVME
^B*S. typhimurium* LAFAAEV^TKNEQMAELL^SG-ALAPET^LAES^FIAV^CGEQ^L--DENG^QN^LIR
^B*V. cholerae* LAFTA^EVAKNEQMH^LLTS-SGSANK^LAE^IFIAV^CGEQ^L--DGHG^QN^LIK
^B*V. alginolyticus* LSFAAEVAKNEQMN^LLTG-SVSAD^KMAE^IFVAV^CGEQ^V--DTHG^QN^LLK
^B*H. influenzae* LGFAAAVAE^DETVKAYL^SS-SLSAQ^KLAD^TVISIC^GEQ^L--DQY^QN^LIR
^B*P. multocida* LHFSS^EL^IKDE^TM^QT^FLKS-SLSAS^KLAD^TVISIC^GEQ^L--DQY^QN^LIR
^B*P. aeruginosa* LGVLA^VSQDD^TV^RQL^LKE^PQ^LTSSA^KAQ^SLID^VCG^DKL--NAPA^QN^FVR
^B*X. fastidiosa* LTFSAQ^VANN^PIAAALL^CH^PQIDHE^QAAALLS---PEGA--DPAY^VR^FLE
^B*T. ferrooxidans* CRRLPAMIAD^VQAQ^AFL^TD^PERRDAE^KV^SLLSAV^PVAVD--VKAW^KAFLA
^B*A. ferrooxidans* CRRLPAMIAD^VQAQ^AFL^TD^PERRDAE^KV^SLLSAV^PVAVD--VKAW^KAFLA
^B*N. meningitidis* Z2491 LEKLA^VVQEG^KVAS^LDR^PETNASE^KAD^LIL^DLV^GLKD--KE-LKN^FVI
^B*R. capsulatus* DALTAALAGS^AELRDL^ISS^VY^TREE^QGDA^IAAVA^AKMG-LSAPLAN^GLK
^B*R. blasticus* DALK^DVL^AGS^PDLGAM^IAS^PVIS^RG^DQAKAVAAI^AG^KMG-LSPL^MT^NLA
^B*A. tumefaciens* DKFGALLDES^DLK^RL^VAS^PV^FSAED^QFKAI^SAICE^KAG-IAGLAV^NFLK
^B*Bacillus* sp. PS3 RAVRQALAE^NGE^FLS^LSY^PK^LSL^DCK^KALIREA--FAG-VST^PV^QNT^LL
^B*B. subtilis* IVVK^QV^FQNEKALND^VL^NH^PK^VPA^KK^KELIQ^NA--FGS-LSQ^SVL^NIT^F
^B*B. firmus* ELVKEV^NTTP^EFL^QLLSH^PK^VTTE^KK^RAFIENS--FK^DSL^SET^SL^HTLL
^B*B. pseudofirmus* ELVKEV^NTTP^EFL^QLLSH^PK^VTTE^KK^RAFIENS--FK^DSL^SET^SL^HTLL
^B*B. halodurans* ELIKAV^FQ^DTPELM^QFL^TH^PK^TEL^TQ^KREL^EKT--FKG^VND^TI^FNT^LV
^B*L. acidophilus* MALKKA^AIAN^PK^FIS^VLS^DIL^SSK^DK^K-SILTA--VEK^DFS^DEV^QG^FLN
^B*E. coli* LAFAAEV^TKNEQMAELL^SG-ALAPET^LAES^FIAV^CGEQ^L--DENG^QN^LIR

Fig. 5. Continued

Mitochondrial OSCP

- ^MYeast
- ^M*S. pombe*
- ^M*K. lactis*
- ^MFruit fly
- ^MRat
- ^MBovine
- ^MHuman
- ^M*A. thaliana*
- ^MSweet potato

```

91      101      111      121      131
VLSENNRLGCFEKIASDFGVLND-AHNGLLKGTVTRAEPRDPKSPFKRIEK
VLLDNHRLYLLTRIQKQFSTLMR-AKRGEIEVKITATPNDKILSRLES
VLAENNELSLLEHVTSQFSLKTD-AYHGLVQATVTRAQPNDSKLFKRVEK
LLADNGQLKKLDTVINAYKTIMA-AHRGEVVCEVVEAKPNDASQSKQLEG
LLAENGLGNTQGVISAFSTIMS-VHRGEVPCTVTRAFPDEAVLSELKT
LLAENGLTNTPAVISAFSTMMS-VHRGEVPCTVTRASANDETTLTELKT
LLAENGLSNTQGVVSAFSTMMS-VHRGEVPCTVTRASPREEATLSELKT
LLAENGLKLNLDIAIVKRFMQLTN-AHRG-VKVIIVTVIIPAPAEKELTE
ILAENGLKHXIDRIVKRFKELTM-AHRGEVKATVTRVIPPAPAEKELKA

```

Bacterial and Chloroplast δ

- ^C*P. purpurea*
- ^C*G. theta*
- ^C*O. neapolitana*
- ^C*Antithamnion* sp.
- ^C*O. sinensis*
- ^C*C. caldarium*
- ^CTobacco
- ^CPea
- ^CSpinach
- ^CSorghum
- ^C*C. reinhardtii*
- ^B*Anabaena* sp. PCC 7120
- ^B*Synechocystis* sp. PCC 6803
- ^B*Synechococcus* sp. PCC 6301
- ^B*M. genitalium*
- ^B*M. gallisepticum*
- ^B*M. thermacetica*
- ^B*P. modestum*
- ^B*C. acetobutylicum*
- ^B*S. bovis*
- ^B*Buchnera* sp. APS
- ^B*R. prowazekii*
- ^B*T. maritima*
- ^B*C. jejuni*
- ^B*H. pylori*
- ^B*S. typhimurium*
- ^B*V. cholerae*
- ^B*V. alginolyticus*
- ^B*H. influenzae*
- ^B*P. multocida*
- ^B*P. aeruginosa*
- ^B*X. fastidiosa*
- ^B*T. ferrooxidans*
- ^B*A. ferrooxidans*
- ^B*N. meningitidis* Z2491
- ^B*R. capsulatus*
- ^B*R. blasticus*
- ^B*A. tumefaciens*
- ^B*Bacillus* sp. PS3
- ^B*B. subtilis*
- ^B*B. firmus*
- ^B*B. pseudofirmus*
- ^B*B. halodurans*
- ^B*L. acidophilus*
- ^B*E. coli*

```

VLVDRKRISMLDVIAGRYLELAY-AMESLTIANISIAIANSQDENLLID
VLCDRGGSANLSSIIDNTIELAY-KKASIEIAYVTRATAFSSNQGEALVE
LLVNRGVAFLDKIAQKFLLELSY-KQDAIEIAKVTRSVASAAQQQKELAG
VLVDRRISMLHLIIDKYFSLVY-KTESTILTEVTRATIDTEEQEIALIN
VLVNRSSINLLEPIIASYLNLY-NAASVKMIEVSRAYAFNLKQNTLIK
ILIERRRIIFYFDIVDQFILIWN-KATNTSVVEIARVISRTEKQQOYLIN
ILVDMKIVELIKEIVKEFEKVYN-TLTDTELAVVTRVVKVESQHLAQIAK
VLIDSKRIDMIIDIIKEFEVYN-TLTDTELVVTRVVKVESHHLAQIAK
ILIDSEINLVKEILNEFEDVFN-KITGTEVAVVTRVVKVENDHLAQIAK
VVVDNFGATILPEIVVEFENIFN-SLTGTEVAVVTRVVKVESQDLAQIAK
LLVEKDELSLVPEICECFEDLYC-QMTDTQVATLRSAVKEQEQQFLIAK
LLVDKRRIFFLPEILQQYLALLR-QLNQTVLAEVTRAVATTEDEQQQAVTE
LLVDKRRIVFLEAICEQYLALLR-QFTNTVLAEVTRALKTDQAKDQVKE
LLVDRNIIAFLDGIADRYQALLR-KLRNVVRADVSRSAVPTAEOVQVITE
VIIILKGYFTLVEQAIKYFFDNVE-SQKHIFQIKIIRAFESSSKQLNKIIA
VVIDNREFFHIKSIKKKFRMIE-KEEHTIFINVVRAHEANDDQKAELE
VLAKGERALPEMAAQFRRLVD-QAENILPVEVTRAITREDILAGLKE
YLIKEDLSSIRGIVAEYKIIYY-AKNQILDVEAIFAIETPKDQKAKLIE
LLLEKKAHDAEGILTQLNKISL-EKHNKVVAEVRVIVPTDNEKTTLAS
VILQNEAEAFKAILEGVQKDFV-IATNQHDIVTRAVATTEDEQKERILA
LLSENQNFNILNNIFERFVKLEA-CYKNIIVQLKGAFSKENLINKIRK
LLIKNSRTHILSNIVEVYNKLLY-ESRNIKIVHVISTNEQQPKEQEWIQS
TLVFERQRQKLLRNIRDLFYEKILSEQVPAANLSIAHSPEDELSLRLK
LLIENSLECIPOIVKELELRQKA-FKENIFVGIYVYKESQENLKDLEV
VVLKNNLDMLELITEELSFDSK---RTLEATLLVPEKNNENELEAVQQ
VMAENGLNALPDVLEQFIHLRA-VSEATAEVDVTRAAAASEQQLAEISA
VMAENGLLAIIPALYERFAVLKQ-EHEKKVDVEVTRATESEQQQRSEIGS
VMAENGLAALPDVCTEFTYLK-EHEKEIDVEVTRATESEDEQLANIGS
LMAENKLSAIPAVFEEFKHVE-EHQAIAEVETSAQPNATQIEKIAA
LMAENKLEVLPAIQAFQHYVE-EHQKVAEVQVTRSAQPNATQEQKIAS
TVAKNKLELLPTIAEMYQLKA-EQEKSVEVEVTRAFTESEKQQDKLAK
VIAEAHLDVLLQVAGLYEKLRA-EAEHVIAKITAIEIAPNELNNIVT
LLIHNDQWPATAEIGTLFADAMR-RAEGVVDVLTBAIAADAGQKTAVQS
LLIHNDQWPATAEIGTLFADAMR-RAEGVVDVLTBAIAADAGQKTAVQS
VLAQKRLSILPEVYAQYQDLTL-SFNHIKSAVIYRAPHNTDKQVGEFAQ
LMATKRRLFALPQLLGLAAAIA-EAKGEMTADVTRATAASAAQAEKLA
LMSKRLFALPQVLSALAGLIA-EKGEVTAETVATKLSAAQAKKLA
VVANNRFLFVPGMIRAYRTIAA-AHRGEITAEVTRAHADDEAQETELKA
LLLEHRHFGVLPELAGTVSRPRSTTARGIAKAVAYGGAASTDEELRALSD
LLIDRHAAIVPELTDEFIKLAN-VARQTEDAIVYVVKPNTDAEMPLPSQ
LLVENKIESLVDMIDSE-KEMSYEAQDMAEAVVYBAKPHNTSEBQAQIAV
LLVENKIESLVDMIDSE-KEMSYEAQDMAEAVVYBAKPHNTSEBQAQIAV
VLVERKIDLIIPVVKQF-KLSYDAQKIAEAFVYBAKPHNTSEAEKDQLSV
LLLEYNRFADLIDIIDQFSLLYD-NENKIASGTATVAVKDDQLERLSE
VMAENGLNALPDVLEQFIHLRA-VSEATAEVDVTRAAAASEQQQAKIASA

```

Fig. 5. Continued

Mitochondrial OSCP

	141	151	161	171	181
^M Yeast	ALSASKLVGGKSLKLENVVKPEIRGGGLIVEIG			---	DKTVDLSISTKIQK
^M <i>S. pombe</i>	RIAKS--KYGKGG--LLVSNKVTPSIIGGLIVEIG			---	DNILDVEVGSRLNN
^M <i>K. lactis</i>	ALAASSLVGGKSLKLENVVKPEIQGGGLIVEVS			---	DRTVDLSTIASRINK
^M Fruit fly	ALKS--FLKGNESLKITSRVDPSIIGGLIVSIG			---	DKYVNMSTIATKVKL
^M Rat	VLNS--FLSKGQILNLEVKTDP SIMGMIVRIG			---	EKYVDMSTAKTKIQK
^M Bovine	VLKS--FLSKGQVLKLEVKDPSIMGMIVRIG			---	EKYVDMSTAKTKIQK
^M Human	VLKS--FLSQGQVLKLEAKTDP SILGMIVRIG			---	EKYVDMSTAKTKIQK
^M <i>A. thaliana</i>	TLQE--IIGAGKKITVEQKIDPSIYGGGLIVEFQ			---	QKVLDMSTIRTRAQQ
^M Sweet potato	TLQE--MVGQGKSVQIEQKIDPTIIGGLVVEFG			---	QKVFDMSTIRTRARQ

Bacterial and Chloroplast δ

^C <i>P. purpurea</i>	KIKAM---TSAKEVKLVISVDPELIGGFTTIQIG			---	SKVIDTSIRGQLKQ
^C <i>G. theta</i>	KLKSM---TSTEQIKLINITVDKTLIGGFKVQIG			---	SKVIDTSIRGQLRQ
^C <i>O. neapolitana</i>	KLKLI---TGAKKIKLALRVEPKLIGGFTVEIG			---	SKLIDTSIRGQLKK
^C <i>Antithamnion</i> sp.	KIKVM---TKGKNVKLITTDQTLIGGFIVRIG			---	SKVIDASLSGKLLK
^C <i>O. sinensis</i>	KLKEL---TNAREIRLVITVSSSLIGGFLIKTN			---	SKVIDFTIKNQLKQ
^C <i>C. caldarium</i>	KLKEI---TKANYIKLKLKIDPSLIGGLIIFKFG			---	SNLIDLSLKGKLLKN
^C Tobacco	GVQRL---TGSKNVRIKTVIDESLVAGFTIRYGN			---	SGSKLIDMSVKKQLED
^C Pea	QVQKL---TGAKKVRTKTLDP SLVAGFTVRYGNT			---	SGKFDMSVKKKLEE
^C Spinach	GVQKI---TGAKNVRIKTVDP SLVAGFTIRYGN			---	EGSKLIDMSVKKQLEE
^C Sorghum	HVQKM---TGAKNVRLKTLDP ELIAGFTVQYGRD			---	SSLIDMSVKKQIEE
^C <i>C. reinhardtii</i>	KLQEL---TGSKNIKLPVIDSSLIAGFVVEYG			---	SSQIDLSVRRGQIER
^B <i>Anabaena</i> sp. PCC 7120	KVLAL---TKARQVELATKVDSDLIIGVVIKVG			---	SQVIDSSIRGQLRR
^B <i>Synechocystis</i> sp. PCC 6803	RVKQL---TGAQAVELETKVDGDIIGVVIKVG			---	SQVFDSSIRGQLRR
^B <i>Synechococcus</i> sp. PCC 6301	KVKQL---TGAAGVEIESQVDADLIIGVVIKVG			---	SQVIDASIRGQLKR
^B <i>M. genitalium</i>	IMEKR---PKTKVYKTEIDRSLSIGIRIESS			---	SHLFDKSNVDELKR
^B <i>M. gallisepticum</i>	KLYKK---FASQVKILYQTDPSLIAGIRIQSN			---	DLIDNSIDGKLLK
^B <i>M. thermoacetica</i>	RLAGITR---RNIRLSSRVNPELIGGVVIRLG			---	DRVIDASVKKKLEL
^B <i>P. modestum</i>	QLEKK---TGKKNVLEVSIDKSIAGGIKIG			---	DEIDGSSVRRQLDT
^B <i>C. acetobutylicum</i>	KLSAK---YNKIIIFKEIIDKTIIGVYVRVG			---	DDVIDGTIKFKLES
^B <i>S. bovis</i>	LVAEK---FGVKAGKLVENIDESIIGGFVINVN			---	NKVIDTSIRRQLQE
^B <i>Buchnera</i> sp. APS	VLERF---FLKKTIIKYVDPNILNGMIVKVN			---	NTIFDLSAQNHLLKQ
^B <i>R. prowazekii</i>	RIEKE---LQHKTELFDDIDNTIIGVVIKYD			---	NVLRDYSIKGSLEK
^B <i>T. maritima</i>	FVRKY---ALKDPVFDISIDESLIAGALVEFEG			---	FRIDTTVQGRLLK
^B <i>C. jejuni</i>	KLNKK---FDANIKNLNKISQDDSVKIELEELG			---	YELSFMSKALQNKLE
^B <i>H. pylori</i>	KLQAR-----FNAPVEITQDTWSKKGVSLSVSS			---	LDLEIGFBSKEDILKK
^B <i>S. typhimurium</i>	AMEKR---LS-RKVKLNCKIDKSMAGVIRAG			---	DMVIDGSSVRRGLER
^B <i>V. cholerae</i>	KLEQR---LE-RKVQLNCSVDETLIGGVIRAG			---	DLVIDNSARGLRKR
^B <i>V. alginolyticus</i>	KLEKR---LE-RKVKLNCSVDETLIGGVIRAG			---	DLVIDDSARGLRNR
^B <i>H. influenzae</i>	AMEKR---LA-RKVKLNCSVDNALIAGVIVRTE			---	DFVIDGSSRGLTR
^B <i>P. multocida</i>	AMEKR---LA-RKVKLNCSLSSLIAGVIRTD			---	DFVIDGSSRGLSR
^B <i>P. aeruginosa</i>	ALSAR---LS-REVRLHASEDASLIIGGVIRAG			---	DLVIDGSSVRRGLAK
^B <i>X. fastidiosa</i>	ALKKR---FD-CEIEVTTGVDHSLIGGAVIDTG			---	NVVIDGSSIRSKLTR
^B <i>T. ferrooxidans</i>	ALERR---FAGHKVRFREAVDAALIGGLVIHTG			---	DLTIDASVRRGQVQQ
^B <i>A. ferrooxidans</i>	ALERR---FAGHKVRFREAVDAALIGGLVIHTG			---	DLTIDASVRRGQVQQ
^B <i>N. meningitidis</i> Z2491	MLNKR---FD-SELKISVEIEPELIGGLIKVEVC			---	DQVLDSVQGLSA
^B <i>R. capsulatus</i>	TLAK----QTGKTVKLNVAVDDESILGMIVKLG			---	SRMIDTTVKAKLAS
^B <i>R. blasticus</i>	TLKA----KVGKTVKLNVTVDDESILGLIVKLG			---	STMIDTSVKSGLAS
^B <i>A. tumefaciens</i>	ALKS----VTGKDVAISVTVDP SILGGLIVKVG			---	SRQIDTSIRTKLST
^B <i>Bacillus</i> sp. PS3	VFAQK---VGKQTFLEIENIIDPELIGGVNVRIG			---	NRIYDGSVSGQLER
^B <i>B. subtilis</i>	VFAKK---AGVASLRIRNEVQTDLIGGIVKVRIG			---	NRIYDGSVSGKLR
^B <i>B. firmus</i>	IFAKK---VNKAKLLVINNVNKDLLGLKIRIG			---	DRIYDGRVKSQDR
^B <i>B. pseudofirmus</i>	IFAKK---VNKAKLLVTNVNKDLLGLKIRIG			---	DRIYDGSVKSQDR
^B <i>B. halodurans</i>	LFPAK---VGKAKLLIENIIVDPSIIGGLKIRIG			---	DRIYDGSIRGQLDV
^B <i>L. acidophilus</i>	SFAKK---YDLNAVRLNKVDP SILGGVILQVK			---	DRVIDGSSVKNKLLK
^B <i>E. coli</i>	AMEKR---LS-RKVKLNCKIDKSMAGVIRAG			---	DMVIDGSSVRRGLER

Fig. 5. Continued

Mitochondrial OSCP

	191
^M Yeast	LNKVLDSI-----
^M <i>S. pombe</i>	LNKLLSEPI-----
^M <i>K. lactis</i>	LNQVLREAI-----
^M Fruit fly	YTDVIQTAA-----
^M Rat	LSKAMRDLL-----
^M Bovine	LSRAMREIL-----
^M Human	LGRAMREIV-----
^M <i>A. thaliana</i>	MERLLREPVDFNLL
^M Sweet-potato	MERFLREPLNF---

Bacterial and Chloroplast δ

^C <i>P. purpurea</i>	MASHLDVAAM----
^C <i>G. theta</i>	LASHLGSSAI-----
^C <i>O. neapolitana</i>	ISNLLGA-----
^C <i>Antithamnion</i> sp.	IAFYLETN-----
^C <i>O. sinensis</i>	LAKHLDSVLEI---
^C <i>C. caldarium</i>	SPYI-----
^C Tobacco	IAAQLEIGDIQLAV
^C Pea	IAAQIDLGDQLAV
^C Spinach	IAAQLEMDDVTLAV
^C Sorghum	ITSEFELPDVPLEV
^C <i>C. reinhardtii</i>	VADQLTKEMTAKLS
^B <i>Anabaena</i> sp. PCC 7120	LSLRLSNS-----
^B <i>Synechocystis</i> sp. PCC 6803	VGLSLGTAL-----
^B <i>Synechococcus</i> sp. PCC 6301	ISISLAA-----
^B <i>M. genitalium</i>	IMAHFI-----
^B <i>M. gallisepticum</i>	LKHQLRTFSKEN--
^B <i>M. thermoacetica</i>	LGEHLKRA-----
^B <i>P. modestum</i>	IARS-----
^B <i>C. acetobutylicum</i>	MKKVMLKEE-----
^B <i>S. bovis</i>	FKMNLK-----
^B <i>Buchnera</i> sp. APS	LSDSLNF-----
^B <i>R. prowazekii</i>	IAKCLKNVKIC---
^B <i>T. maritima</i>	IAREALKRGEMS--
^B <i>C. jejuni</i>	YILKII-----
^B <i>H. pylori</i>	IEKQVIQSI-----
^B <i>S. typhimurium</i>	LADVLQS-----
^B <i>V. cholerae</i>	LSDALQS-----
^B <i>V. alginolyticus</i>	LSDALQS-----
^B <i>H. influenzae</i>	LANELQL-----
^B <i>P. multocida</i>	LANELQL-----
^B <i>P. aeruginosa</i>	LAEALKS-----
^B <i>X. fastidiosa</i>	LQASLTH-----
^B <i>T. ferrooxidans</i>	LARTLRS-----
^B <i>A. ferrooxidans</i>	LARTLRS-----
^B <i>N. meningitidis</i> Z2491	LYTTMTN-----
^B <i>R. capsulatus</i>	LQNAMKEVG-----
^B <i>R. blasticus</i>	LQNAMKEVG-----
^B <i>A. tumefaciens</i>	LKLALKEVG-----
^B <i>Bacillus</i> sp. PS3	IRRQLIG-----
^B <i>B. subtilis</i>	IERQLAGENR----
^B <i>B. firmus</i>	LERQLIAGTR----
^B <i>B. pseudofirmus</i>	LERQLIAGTR----
^B <i>B. halodurans</i>	LHRELVSGPRS---
^B <i>L. acidophilus</i>	IRAQIIDEN-----
^B <i>E. coli</i>	LADVLQS-----

171

Fig. 5. Continued

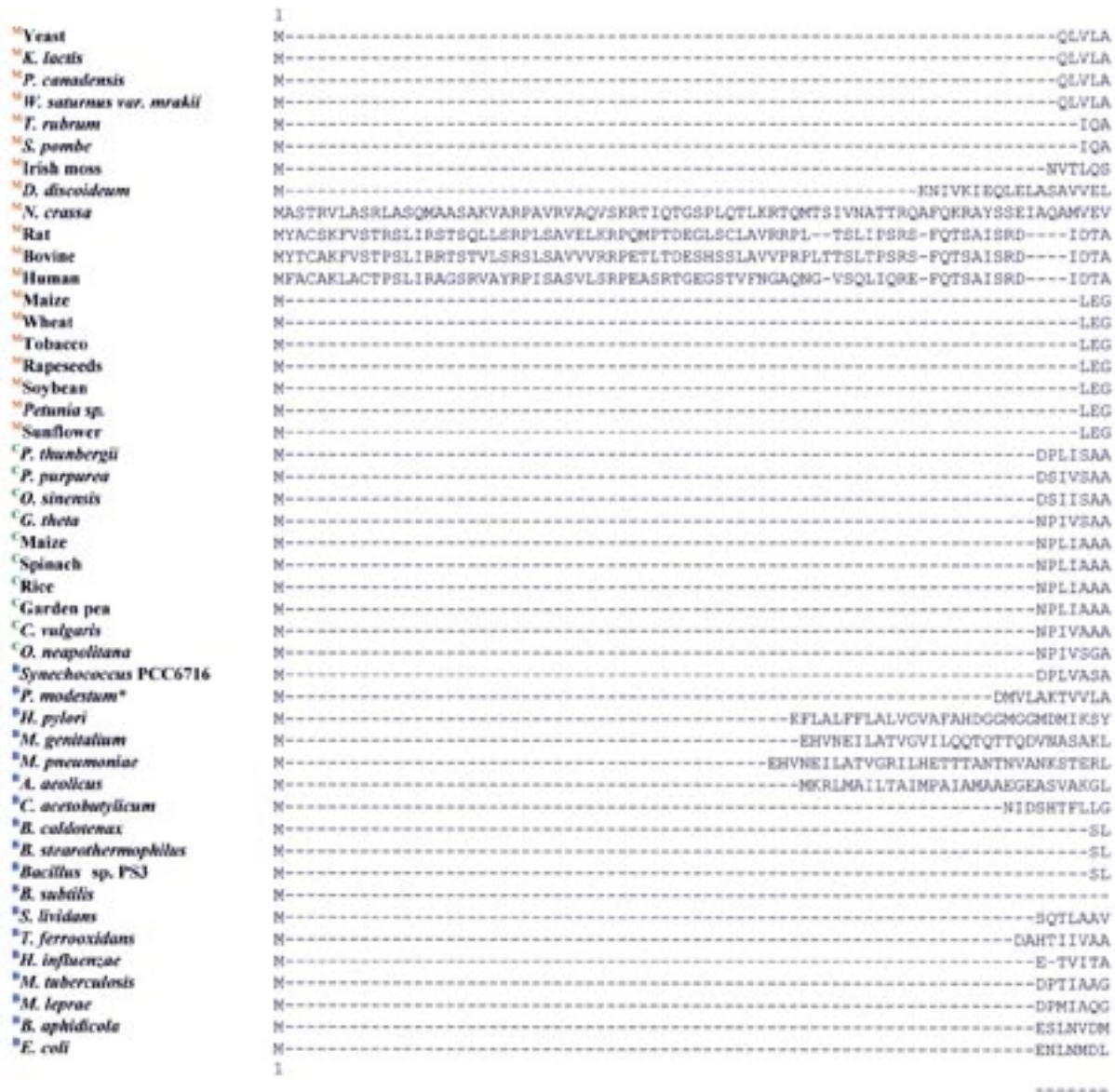


Fig. 6. Sequence alignment of subunit c. Completely, highly, and weakly conserved residues are highlighted in bold and shaded in red, purple, and yellow, respectively. The TM1 and TM2 below the sequences represent transmembrane helices. The interhelical "loop" region is indicated by a red dotted underline.

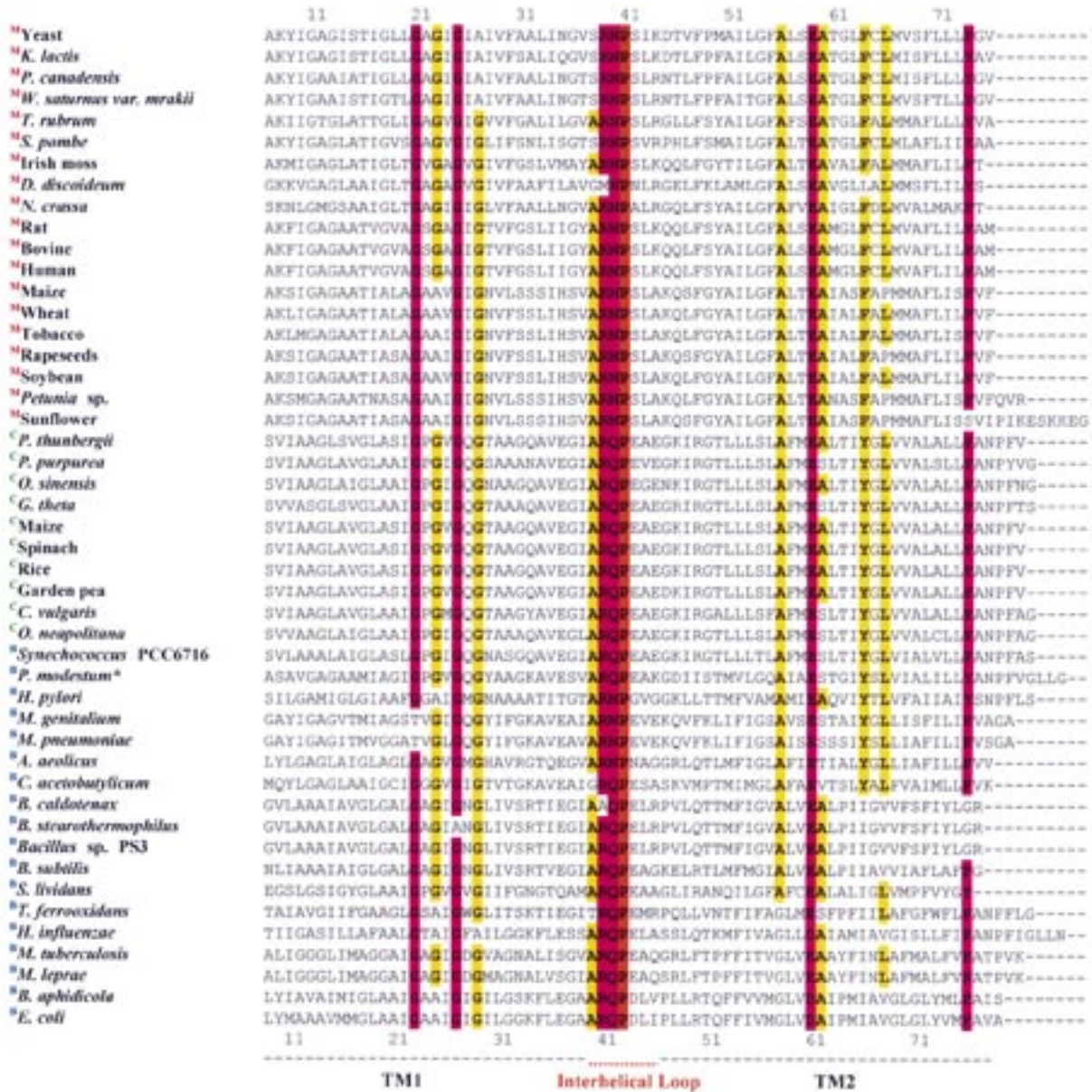


Fig. 6. Continued

Length and Dimerization Domain of Mitochondrial Subunit B

The mature mitochondrial subunit *b* is approximately 1.2–1.4 times longer than that of bacterial subunit *b*. Bacterial subunit *b* forms a dimer at its C-terminal domain, and in recent studies using the *E. coli* protein, residues 53–122 were reported to form this dimer, which can span 100–110 Å (Dunn *et al.*, 2000; Revington *et al.*, 1999, 2002). In mitochondrial subunit *b*, no studies have been performed

on the dimerization domain. To predict the dimerization domain of mitochondrial subunit *b*, we aligned and compared sequences of subunits from bacteria, chloroplasts, and mitochondria, and then used the coiled-coil prediction algorithm, which is based on the relative frequency of occurrence of amino acids at each position of the coiled-coil heptad repeat (Lupas *et al.*, 1991).

In the sequence alignment of subunit *b*, no significant homologies are found between mitochondria and bacteria, and between mitochondria and chloroplasts (Data not

δ subunit from mitochondria

		1
^Y Yeast	M---LRSIIGKSASR-----	SLNFVAKRSYAEAAAA
^A <i>A. bisporus</i>	M--SSLRLLAS-----	AARRATTHVAYTR
^N <i>N. crassa</i>	M-----	
^K <i>K. lactis</i>	M---FRLSAAR-----	TLAKSVNTVVAKR
^B Bovine	M---LPSALLRRPGL-----	GRLVRQVRLYAEAAAA
^H Human	M---LPAALLRRPGL-----	GRLVRHARAYAEAAAA
^R Rat	M---LPAALLRH PGL-----	RRLVLQARTYAQAAA
^C <i>C. elegans</i>	M---LARTIQR-----	FSVVAKRGYAAAA
^S Sweet potato	MFRHSSRLRLARATTMGWRRPFSTDLPAETAADSTFVEAWKKLIPNVDPPKTPSAYMAPRP	

 ϵ subunit from bacteria and chloroplasts

^A <i>A. columnaris</i>	M-----	
^W Wheat	M-----	
^B Barley	M-----	
^M Maize	M-----	
^R Rice	M-----	
^A <i>A. thaliana</i>	M-----	
^S Spinach	M-----	
^S Sweet potato	M-----	
^T Tobacco	M-----	
^A Alfalfa	M-----	
^P Pea	M-----	
^N Norway spruce	M-----	
^J Japanese black pine	M-----	
^L Liverwort	M-----	
^C <i>C. reinhardtii</i>	M-----	
^C <i>C. vulgaris</i>	M-----	
^E <i>E. gracilis</i>	M-----	
^C <i>C. caldarium</i>	M-----	
^P <i>P. purpurea</i>	M-----	
^O <i>O. sinensis</i>	M-----	
^P <i>P. littoralis</i>	M-----	
^S <i>Synechocystis</i> PCC6803	M-----	
^A <i>Anabaena</i> PCC7120	M-----	
^S <i>Synechococcus</i> PCC6716	M-----	
^P <i>P. dlademni</i>	M-----	
^B <i>Bacillus</i> sp. PS3	M-----	
^B <i>B. stearothermophilus</i>	M-----	
^B <i>B. caldotenax</i>	M-----	
^B <i>B. megaterium</i>	M-----	
^B <i>B. subtilis</i>	M-----	
^B <i>B. firmus</i>	M-----	
^E <i>E. hirae</i>	M-----	
^S <i>S. mutans</i>	M-----	
^C <i>C. acetobutylicum</i>	M-----	
^M <i>M. thermoacetica</i>	M-----	
^M <i>M. leprae</i>	M-----	
^M <i>M. tuberculosis</i>	M-----	
^S <i>S. lividans</i>	M-----	
^V <i>V. alginolyticus</i>	M-----	
^H <i>H. influenzae</i> Rd	M-----	
^B <i>B. aphidicola</i>	M-----	
^T <i>T. ferrooxidans</i>	M-----	
^R <i>R. blasticus</i>	M-----	
^R <i>R. capsulatus</i>	M-----	
^R <i>R. rubrum</i>	M-----	
^A <i>A. aeolicus</i>	M-----	
^H <i>H. pylori</i>	M-----	
^R <i>R. prowazekii</i>	M-----	
^E <i>E. coli</i>	M-----	

Fig. 7. Sequence alignment of subunits δ (mitochondria) and ϵ (bacteria and chloroplasts). Conserved residues are highlighted the same as in Fig. 6

δ subunit from mitochondria

	11	21	31	41	51
^Y Yeast	-----SSGLKLFALPHETLSSGSEVTQVNLPAKSERIIVLANHVPVTE-QLLPGVVEV				
^A <i>A. bisporus</i>	RGYAEISDKLKLSLALPHKAISSQDVVQVNI PAESQDMHILSSHVPSIE-PLRPGVVEV				
^N <i>N. crassa</i>	-YAEAVADKIKLSLSLPHQAIYKSDVVQVNI PAVSSEMIVLANHVPVTE-QLKPGVVEV				
^K <i>K. lactis</i>	TYAEAADGALKLQFALPHQTLSSGTFVTVQVNLPAKSSQIIVLANHVPVTE-QLVPGVVEV				
^B Bovine	QAPAAAGPGQMSFTFASPTQVFFNSANVRQVDVPTQTGAFHILAAHVPVTLQ-VLRPGLVVV				
^H Human	PAAASGPNQMSFTFASPTQVFFNGANVRQVDVPTLTGAFHILAAHVPVTLQ-VLRPGLVVV				
^R Rat	PAPAAGPGQMSFTFASPTQVFFDGANVRQVDVPTLTGAFHILASHVPVTLQ-VLRPGLVVV				
^C <i>C. elegans</i>	PAANANPEELRLTFASPTQVAVSNVVKQVDVPTLGMVSVLANHVPVTEIG-VLKPGLVSV				
^S Sweet potato	ATPSSIIPSKLTVNFVLPYSSELAGKEVDMVII PATTQMGVLPGLVAVATIA-ELKPGVMSV				

ε subunit from bacteria and chloroplasts

^A <i>A. columnaris</i>	-----KLNLYVLTTPKRIIWD-CEVKEIILSTNSQIIVLPHNAPINT-AVDMGFLRI					
^W Wheat	-----KLNLYVLTTPKRIIWD-CEVKEIILSTNSQIIVLPHNAPINT-AVDMGFLRI					
^B Barley	-----KLNLYVLTTPKRIIWD-CEVKEIILSTNSQIIVLPHNAPINT-AVDMGFLRI					
^M Maize	-----KLNLYVLTTPKRIIWD-CEVKEIILSTNSQIIVLPHNAPINT-AVDMGFLRI					
^R Rice	-----KLNLYVLTTPKRIIWD-CEVKEIILSTNSQIIVLPHNAPINT-AVDMGFLRI					
^A <i>A. thaliana</i>	-----TLNLCVLTTPNRIVWD-SEVEEIIILSTNSQIIVLANHAPVAT-AVDIGILTI					
^S Spinach	-----TLNLCVLTTPNRSIWN-SEVKEIILSTNSQIIVLPHNAPVAT-AVDIGILRI					
^S Sweet potato	-----TLKLCVLTTPNRIVWD-SEVKEIILSTNSQIIVLPHNAPVAT-AVDIGILRI					
^T Tobacco	-----TLNLSVLTTPNRIVWD-SEVEEIVLSTNSQIIVLPHNAPVAT-AVDIGILRI					
^A Alfalfa	-----TLNLCVLTTPNRIVWD-SEVKEIILSTNSQIIVLPHNAPVAT-AVDIGILKI					
^P Pea	-----TFNLCVLTTPNRIVWD-SEVKEIILSTNSQIIVLPHNAPVAT-AVDIGILRI					
^N Norway spruce	-----TLNLRVLTSPNRIVWD-SEVKEIILSTNSQIIVLPHNAPVAT-AVDIGVMI					
^J Japanese black pine	-----TLNLRVLTSPNRIVWD-SEVKEIILSTNSQIIVLPHNAPVAT-AVDIGVMI					
^L Liverwort	-----LNLRIIAPDRIVWD-SDIQEIIILSTNSQIIVLPHNAPVAT-AVDIGIVKI					
^C <i>C. reinhardtii</i>	-----SLQISILTPERPFYN-GQADEIILPTETGEMVVLKNNAPVAT-GLNVGAMLI					
^C <i>C. vulgaris</i>	-----TLQVCIMTPDRIFYN-DQADEIILPTNTQMGVLTNNAPVAT-ALDVGITLI					
^E <i>E. gracilis</i>	-----TLQVSIITIPERVFWE-KRVEEIILPTLQMGVLTNDHPIIPLT-GLDIGIILV					
^C <i>C. caldarium</i>	-----GLNIRVIAPDRIVWN-AKAEVILPTSTQQLSILSGHAPVAT-ALDVGIMRV					
^P <i>P. purpurea</i>	-----TLNIRIAPDRIVWD-AEAQEIILPSTTQQLSILSGHAPVAT-ALDVGIMRV					
^O <i>O. sinensis</i>	-----VMNVRVLTTPTRVICS-TTAEDEVILPGLTGLVHILDGHAALIT-ALDTGLLRI					
^P <i>P. littoralis</i>	-----SLNIRVIAPDGLIWD-TTAEGVVLPSTTQQLSILSGHAPVAT-SLEIGILRI					
^S <i>Synechocystis</i> PCC6803	-----TLTVRVITPDKVVD-EEVQELILPSTTQQLSILSNHAPVAT-ALEIGVIMRV					
^A <i>Anabaena</i> PCC7120	-----TLTVRVITPDKTVWD-AEVDEVILPSTTQQLSILSGHAPVAT-ALDVGILRV					
^S <i>Synechococcus</i> PCC6716	-----VMTVRVIAPDKTVWD-APAEVILPSTTQQLSILSNHAPVAT-ALETGVIMRV					
^P <i>P. didemni</i>	-----TLTLRVITPDKTVWD-DSVEEIVLPSTTQVTVLTGHAALIT-ALDTGVIMRV					
^B <i>Bacillus</i> sp. PS3	-----KTIHVSVVTPDGPVYE-DDVEMVSVKAKSSELGILPGHPIIPLVA-PLIISAARL					
^B <i>B. stearothermophilus</i>	-----KTIHVSVVTPDGPVYE-DDVEMVSVKAKSSELGILPGHPIIPLVA-PLIISAARL					
^B <i>B. caldotenax</i>	-----KTIHVSVVTPDGPVYE-DDVEMVSVKAKSSELGILPGHPIIPLVA-PLIISAARL					
^B <i>B. megaterium</i>	-----KTIHVSVVTPDGPVYE-SEVEMVSTRAQSSELGILHGHPIIPLVA-PLQIGAVRL					
^B <i>B. subtilis</i>	-----KTVKVNIVTPDGPVWD-ADIEMVSVRAESSELGILPGHPIIPLVA-PLKIGAVRL					
^B <i>B. firmus</i>	-----STIRVNVVTPDGPVWD-GVDVLLVVVTVENELGIFRKHPIIPLVA-PVTVGRVRL					
^E <i>E. hirae</i>	-----DQYLTVNVVTPDGLVWD-HHAAIVVARTTAGEIILPKHPIIIV-PLTIDEVVR					
^S <i>S. mutans</i>	-----AEMTVQIVTPDGLKWD-HHAKFIVKTPNSELVLANHENLIA-PLEVHEMKI					
^C <i>C. acetobutylicum</i>	-----ANNIKLSILTPQKTFV-GDVKEIITRTVEEIIILPNHPIIPLVA-PLTPTETIL					
^M <i>M. thermoacetica</i>	-----ASLNLEIITPERVVVQ-AEAASVIAPGICSYLIVLPEHAPVAT-PLQAGVVVTC					
^M <i>M. leprae</i>	-----LCMDLNIEIVAVDRKIWS-GKGTFLFTRTTAGEIILPRHPIIPLVAQLVDDANVRI					
^M <i>M. tuberculosis</i>	-----AELNVEIVAVDRNIWS-GTAKFLFTRTTVEEIIILPRHPIIPLVAQLVDDANVRI					
^S <i>S. lividans</i>	-----AELHVALVAADREVWS-GEATLVVARTTSDIIVMFGHPIIPLVA-VLESQPVTI					
^V <i>V. alginolyticus</i>	-----AATFHLDVVSAEKIWS-GRVETFQVTGSESELGIFHGHTPLIT-AIKPGMVRI					
^H <i>H. influenzae</i> Rd	-----ATFNLTIVSAEQIWE-GEVQIQVTGVESELGILPGHTPLIT-AIKPGIVKF					
^B <i>B. aphidicola</i>	-----DFYLDIVSVEKRIWS-GLVDRIQVSGSESEMIIYPGHTQLLS-IIKPGVIYI					
^T <i>T. ferrooxidans</i>	-----AMTIDVRVVS AEGSIYA-GVADMVAPGEMSELGILPRHPIIPLVA-GLRPGELRI					
^R <i>R. blasticus</i>	-----AATLQFDLVS PERRLAS-VQATEVQIPGAAVDMTAMQGHAPTIT-TLRPGILRA					
^R <i>R. capsulatus</i>	-----ADTMQFDLVS PERRLAS-VAASEVRLPGVEDLTAMPGHAPVIL-SLRPGIITV					
^R <i>R. rubrum</i>	-----AETTEFELVSPERLLS-EPVEMVVVPGTDGDFAMPRHAPVILS-TVRPGVIST					
^A <i>A. aeolicus</i>	-----IQVEIVSPQMVYS-GEVESVNVPTVEVEVILENHYMLT-LLKPGLVYF					
^H <i>H. pylori</i>	-----MALLKISVVVPEGEVYT-GEVKSVVLPVGEVEFVLYGHSNMIT-LLQAGVIEI					
^R <i>R. prowazekii</i>	-----HETIRVKIITPSSIAE-KQSKMVTMPGEDMFFVLPVHIVMIV-NLKAGLVQI					
^E <i>E. coli</i>	-----AMTYHLDVVS AEQQMS-GLVEKIQVTGSESELGILPGHAPVAT-AIKPGMIRI					
	1	11	21	31	41	51

Fig. 7. Continued

δ subunit from mitochondria

	61	71	81	91	101	111
¹⁶ Yeast	M-EG----	SNSKFFISG	FATVQ	PDSQLCVTAIE	APPLESFSQENIKNLLAEAKKNVS	
¹⁶ <i>A. hisporus</i>	VEDS----	GSQK-WFVSG	FATVHPN	NRLTINVVE	AAPLED	FSIEATRANLQEAANKVAA
¹⁶ <i>N. crassa</i>	IEES----	GSNKQYFLSG	FAVVQ	PGSKLSINAVE	GYALEDFSAEAVRAQIAEAQKIVS	
¹⁶ <i>K. lactis</i>	LE-----	GSKFFVSG	FATVQ	PDSTLAI	TVEAPPLES	FSFENVRSLLAEAQKNVS
¹⁶ Bovine	HAED----	GTTSKYFVSS	SVTVNADSSVQ	LAAEAVTLDMLD	LGAAKANLEKAQSELL	
¹⁶ Human	HAED----	GTTSKYFVSS	SIAVNADSSVQ	LAAEAVTLDMLD	LGAAKANLEKAQAEVL	
¹⁶ Rat	HAED----	GTTTKYFVSS	SVTVNADSSVQ	LAAEAVTLDMLD	LGAAKANLEKAQSELS	
¹⁶ <i>C. elegans</i>	TTNE----	GTVQRLFVSS	TVSVNIDGSCQV	LAAEVLKVEE	IDESARAELDAQRASG	
¹⁶ Sweet potato	HE-G----	NDVSKYFVSG	FATIHANSFADI	IAVEAVPLDRI	DANLVQKGLAEFTQKLN	

ε subunit from bacteria and chloroplasts

¹⁶ <i>A. columnaris</i>	RLLN----	DQWLTAVLWS	FARIVNNE-I	IILGND	AELGSDID	PEEAQKALEIAEANLS								
¹⁶ Wheat	RLLN----	DQWLTAVLWS	FARIVNNE-I	IILGND	AELGSDID	PEEAQKALEIAEANLS								
¹⁶ Barley	RLLN----	DQWLTAVLWS	FARIVNNE-I	IILGND	AELGSDID	PEEAQKALEIAEANLS								
¹⁶ Maize	RLLN----	DQWLTAVLWS	FARIVNNE-I	IILGND	AELGSDID	PEEAQKALEIAEANLS								
¹⁶ Rice	RLLN----	DQWLTAVLWS	FARIVNNE-I	IILGND	AELGSDID	PEEAQKALEIAEANVS								
¹⁶ <i>A. thaliana</i>	RLA-----	NQWLTALMG	FARIGNNE-IT	ILVND	AEKNSDID	PQEAQOTLEIAEANLR								
¹⁶ Spinach	RLN-----	DQWLTALMG	FARIGNNE-IT	ILVND	AERGSID	PQEAQOTLEIAEANLR								
¹⁶ Sweet potato	RLN-----	DQWVTALMG	FARIGNNE-IT	VLVND	AEKGSID	SLEAQOTLEIAEANFR								
¹⁶ Tobacco	RLN-----	DQWLTALMG	FARIGNNE-IT	VLVND	AEKGSID	PQEAQOTLEIAEANVK								
¹⁶ Alfalfa	RLTN-----	QWVTALMG	FARIGNNE-IT	ILVND	AEKSID	PQEAQOTLEIAEANLN								
¹⁶ Pea	RLK-----	DRWLTALMG	FARIGNNE-IT	ILVTD	AEASDIN	PQEAQOTLEIAEANLN								
¹⁶ Norway spruce	RLN-----	GQWSTALMG	FAKIDNDR-IT	VLVND	AEKSDID	DLKEAQETFKVAKADLA								
¹⁶ Japanese black pine	RLN-----	GQWSTAMMG	FAKIDSDR-IT	VLVND	AEKSDID	PREAQENFRIAKADLA								
¹⁶ Liverwort	RLN-----	DQWSTALMG	FAMIDNNN-LT	ILVND	AEKASEID	YQEAQETFKQAKTNLE								
¹⁶ <i>C. reinhardtii</i>	RGQASGSK	DEWNSYALMG	FALVKQNC-VT	ILVND	AEVSAENIN	PEEAQDAFETAKANLE								
¹⁶ <i>C. vulgaris</i>	R-----	SNSNWNFVALMG	FALVKQNC-VT	ILVND	AEASAQ	IGVDEAEIAPQEAQTKLE								
¹⁶ <i>E. gracilis</i>	R----	QKSSDWTSLVVT	CFALINSNN-VT	ILVND	AEFGSEIN	VEQAQISYNSSKHALE								
¹⁶ <i>C. caldarium</i>	RITN-----	TWTSIVLFG	FAEVENDE-IL	ILVND	AEAEASVINL	DKANKELIESSLLN								
¹⁶ <i>P. purpurea</i>	RVDK-----	EWPIVLLG	FAEVENNQ-LT	ILVND	AEAEASQID	LVEAEKNDLDTATQLLN								
¹⁶ <i>O. sinensis</i>	KLN-----	EKWTPILCG	GLAEIDRNR-VT	VLVND	VEELVAVELNE	ATTELEKATFLAVE								
¹⁶ <i>P. littoralis</i>	KTN-----	SKWTPILV	GFVAVIKDDE-VL	VLI	SGVEEV	IKEDYSFAKSI	LAKAKIDLD							
¹⁶ <i>Synechocystis</i> PCC6803	RP GK-----	DWQNI	AVMGGFAEVENNE-VK	VLVND	AELEGGT	IDAESARQAYTAAQGALE								
¹⁶ <i>Anabaena</i> PCC7120	RTSKS-----	QNQAI	IALLGFAEVEEDE-VT	ILVND	GGERGDT	INLEEAR	TAYSQAQTKLN							
¹⁶ <i>Synechococcus</i> PCC6716	RQER-----	EWVAI	ALMGFAEVENNE-VT	VLVND	AEERGD	IDLETAKREFSEAAVA								
¹⁶ <i>P. didemni</i>	RP GK-----	DWQAI	ALMGFAEVENNE-VK	VLVND	AEVGDSD	DKETARTEPQAAQNL								
¹⁶ <i>Bacillus</i> sp. PS3	KKGGKT-----	QYIAV	SGGFLEVRPDN-VT	ILAQAA	AERAED	IDVLR	AKARKS-GRTPFQ							
¹⁶ <i>B. stearothermophilus</i>	KKGGKT-----	QYIAF	SGGFLEVRPDN-VT	ILAQAA	AERAED	IDVLR	AKARKS-GRTPFQ							
¹⁶ <i>B. caldotenax</i>	KKGGKT-----	QYIAV	SGGFLEVRPDK-VT	ILAQAA	AERAED	IDVLR	AKARKS-GRTPFQ							
¹⁶ <i>B. megaterium</i>	KKASST-----	ELVAV	SGGFLEVRPDK-VT	ILAQAA	ETAEE	IDVAP	AEAKKRAEMRLD							
¹⁶ <i>B. subtilis</i>	KKDQQT-----	EMVAV	SGGFLEVRPDH-VT	ILAQAA	ETAEG	DKER	AEAAQRAQERLN							
¹⁶ <i>B. firmus</i>	KKGNSE-----	EQVAV	SGGFLEVRADQ-VT	ILAEAA	ELPSA	IDVDR	ARAAKERAESRLN							
¹⁶ <i>E. hirae</i>	KRTDSDTH---	VDWIAV	NGGIMEVRDNV-VS	IVADS	AERERD	IDVSR	AERAKQRAERQIA							
¹⁶ <i>S. mutans</i>	KRIDDDSH---	VDWVAV	NGGIIIEIKDNL-VT	IVADS	AERERD	IDL	SRAERAKKRAEKAIE							
¹⁶ <i>C. acetobutylicum</i>	VEEDGSR----	KKVFT	STILNVGESE-VS	FMCD	ASEWPDE	IDIQRA	ETAKERAEKRLK							
¹⁶ <i>M. thermoacetica</i>	RRRERA----	EERVAV	SGGFLEAGPDQ-VI	ILADT	AERSEE	IDVEM	ARQARERAEERLR							
¹⁶ <i>M. leprae</i>	EREGEKD----	LRVAV	DGFFLSVTEER-VS	ILAE	SAEFDSE	IDENA	AK-----QDAES							
¹⁶ <i>M. tuberculosis</i>	EREGEKD----	LRIAV	DGFFLSVTEEG-VS	ILAE	SAEFESE	IDEA	AK-----QDS							
¹⁶ <i>S. lividans</i>	RTSDGGT----	VVA	AVHGGFISFADNK-LS	LLAE	VAELSDE	IDVH	FAERKLE--QAKTE							
¹⁶ <i>V. alginolyticus</i>	VKQH----	GHEE	IIVVSGGHEVIQ	FGT-AT	VLADT	AIRGEEL	DAAKAEAKRRAEEQIQ							
¹⁶ <i>H. influenzae</i> Rd	TLKD----	GNEE	VIYVSGGFLEVQ	PNI-VT	VLADT	AIRGSEL	DADRIHEAKRKAENIV							
¹⁶ <i>B. aphidicola</i>	FHKH----	KTEE	CLYISGFI	LEVQPSV-VS	ILAD	VAIRGIDL	DRKRVV	KAKKQAEYFK						
¹⁶ <i>T. ferrooxidans</i>	IHG-----	AETE	FLVNGI	LEIQPDM-VT	VLADS	AERATD	IDEAF	ALAAQAAERMA						
¹⁶ <i>R. blasticus</i>	VS-A-----	EGTK	AVVTGFAE	ISATG-VS	VLAE	AVPLDEM-DA	KIMDQ	LVADASAAS						
¹⁶ <i>R. capsulatus</i>	VS-A-----	AGTAE	YAVTGGFAE	VSGEK-VT	VLAE	RGLTRAEL-TA	AVHAEM	LAEAKKVA						
¹⁶ <i>R. rubrum</i>	YNGG-----	KVQR	RI	FVAGFAE	VTEDR-CT	VLADE	AFD	LASLSEAVRARLQAADDRLK						
¹⁶ <i>A. aeolicus</i>	NGDD-----	K-NG	IAVTV	IVL	VPQKVL	ILAE	EAYEV	GKLP	PASKLKEE	FEAVKMA				
¹⁶ <i>H. pylori</i>	ETEN-----	QKEH	IAIN	GYAE	VTNER-VD	ILAD	CAV	FIKKE	SDRDD-ATS	RAKLL				
¹⁶ <i>R. prowazekii</i>	YIYN---	IHN	YENT	YLISG	VT	EITSHY-IN	IVTE	VA	INVT	NLSESE	ISTQ	RYELQKLLS		
¹⁶ <i>E. coli</i>	VKQH-----	GHEE	FIYLSG	FI	LEVQ	PGN-VT	VLADT	AIR	QDID	EAR	A	MEAKR	RAE	EHIS

Fig. 7. Continued

δ subunit from mitochondria

	121	131
¹⁰ Yeast	S---SDAREA- AEAAIQ VEVLENLQSVLK-----	
¹¹ <i>A. bisporus</i>	G---SGSEADK MEAQI EAEVYEALQHALAK-----	
¹² <i>N. crassa</i>	G---GGSQQDIA EAQ VELEVLESQAVLK-----	
¹³ <i>K. lactis</i>	----SADEVAA EA AIQLEVLLEALQAALK-----	
¹⁴ Bovine	G---AADEATRA EIQIRI EANEALVKALE-----	
¹⁵ Human	G---TADEATRA EIQIRI EANEALVKALE-----	
¹⁶ Rat	G---AADEAARA EIQIRI EANEALVKALEASHVPTLQVL	
¹⁷ <i>C. elegans</i>	----EGSEVAR EAQIRAEV AEALKAATNQ-----	
¹⁸ Sweet potato	T---ASTDVEKA EAQIG VDVHSALNAALTG-----	

ε subunit from bacteria and chloroplasts

¹⁹ <i>A. columnaris</i>	KAE--G T KDL- VEAKL ALRRARIRIEAVNWIPPSN----	
²⁰ Wheat	KAE--G T KDL- VEAKL ALRRARIRIEAVNWIPPSN----	
²¹ Barley	KAE--G T KSL- VEAKL ALRRARIRIEAVNWIPPSN----	
²² Maize	KAE--G T KEL- VEAKL ALRRARIRIEAVNWIPPSN----	
²³ Rice	RAE--G T KEL- VEAKV ALRRARIRIEAVNWIPPSN----	
²⁴ <i>A. thaliana</i>	KAE--G K RQT- IEANL ALRRARTRVEALNTI-----	
²⁵ Spinach	KAE--G K RQK- IEANL ALRRARTRVEASNTISS-----	
²⁶ Sweet potato	KAE--G K RQT- IEANL ALRRARTRVEAINAIS-----	
²⁷ Tobacco	KAE--G R RQK- IEANL ALRRARTRVEAINPIS-----	
²⁸ Alfalfa	KAE--G K RQT- IEANL ALRRARTRVETILE SINRF-----	
²⁹ Pea	KAE--G K RET- IEANL SLRRAKTRVEAIVETIKRIS----	
³⁰ Norway spruce	RAE--G K RQA- IEADV ALKRARTRLEAISASPPVSN----	
³¹ Japanese black pine	RAE--G K RQA- IEADL ALKRARTRLEAINASPPVSN----	
³² Liverwort	EAE--G N KKKE IEALL VFKRAKARLEAINMASKL-----	
³³ <i>C. reinhardtii</i>	KAE--G V KEK- VEANF AYKRAKARYQVVKVLKK-----	
³⁴ <i>C. vulgaris</i>	QSQ--G E KQR- VEATF VFKRARARYQVVKQLGV-----	
³⁵ <i>E. gracilis</i>	MNK--D I KRK- FELTL NLKKARARFQVTLQKK-----	
³⁶ <i>C. caldarium</i>	EAK---T N KE FEATQ KLKAKARVQAANTLTNQSIY----	
³⁷ <i>P. purpurea</i>	DAS---S S KE IEATQ IRKARARVQAATAASA-----	
³⁸ <i>O. sinensis</i>	NAE---T S KAR LDAS IELKKAVARLEGMNYLS-----	
³⁹ <i>P. littoralis</i>	SAK---T T KE IIDA SQELKIASAKVKAFKFI-----	
⁴⁰ <i>Synechocystis</i> PCC6803	EANR GED KPNQLK ASNNY KARARLQAAGGAV-----	
⁴¹ <i>Anabaena</i> PCC7120	QVP-AG DRQAQIQANQA FKRARARFQATGGLA-----	
⁴² <i>Synechococcus</i> PCC6716	KAAQ SGSKQAQIQAAQA FRRARARLQAAGGVVEI-----	
⁴³ <i>P. didemni</i>	RANQ GDNRQELIQATQ EFKRARARFQAAGGMT-----	
⁴⁴ <i>Bacillus</i> sp. PS3	S---Q Q DD IDFKRA ELALKRAMNRLSVAEMK-----	
⁴⁵ <i>B. stearothermophilus</i>	S---Q Q DD IDFKRA ELALKRAMNRLSVAEMK-----	
⁴⁶ <i>B. caldotenax</i>	S---Q Q DD IDFKRA ELALKRAMNRLSVAEMK-----	
⁴⁷ <i>B. megaterium</i>	S---Q Q DD VVKRAE IALKRAVNRLDISQRF-----	
⁴⁸ <i>B. subtilis</i>	S---Q S DD DIRRA ELALQRALNRLDVAGK-----	
⁴⁹ <i>B. firmus</i>	ST--Q Q DA VDFKRA ELALKRAINRLDVTGK-----	
⁵⁰ <i>E. hirae</i>	EAKE KEDT NELK RATVAL HRAINRIKVSKHS-----	
⁵¹ <i>S. mutans</i>	EAKE QHRIDEVQRA QVALRRALNRINVGSK-----	
⁵² <i>C. acetobutylicum</i>	T---S N ND VKRAE LSLSRALARIKTND-----	
⁵³ <i>M. thermoacetica</i>	E---R P PGL DVARAE ALRRARVARLKAAGAI-----	
⁵⁴ <i>M. leprae</i>	D-----D P RIA ARGRARL RAVGAI-----	
⁵⁵ <i>M. tuberculosis</i>	D-----D P RIA ARGRARL RAVGAI-----	
⁵⁶ <i>S. lividans</i>	G-----D A HA ERRAD VRLRAAAGR-----	
⁵⁷ <i>V. alginolyticus</i>	NQH---G D MD FAQA ASELAKAIAQLRVIELTKKRR----	
⁵⁸ <i>H. influenzae</i> Rd	SRG-SD ADH DL LVAKL SKELAKLRAYELTEKLLKTRR--	
⁵⁹ <i>B. aphidicola</i>	KET---T N V KKDDV LLEISKAI AKLRV LEIMDKFKK---	
⁶⁰ <i>T. ferrooxidans</i>	GHT-D Q ME YAAAQ ELLE QIAR LKTVQRLREQGFVR---	
⁶¹ <i>R. blasticus</i>	S---V G -- V DK DTAE KAMSDLQAMKAAAGF-----	
⁶² <i>R. capsulatus</i>	D---A A HP SVADA AK MLAD MEALGSHINL-----	
⁶³ <i>R. rubrum</i>	E---A T SE AKEAE AQAKA IAE ALLAARKG-----	
⁶⁴ <i>A. aeolicus</i>	T---A Q T MEEL KWEKE AEKAR TLLLELVEKYR-----	
⁶⁵ <i>H. pylori</i>	DAS-----S D RL AVSSV LAKIESL-----	
⁶⁶ <i>R. prowazekii</i>	HQH-----	
⁶⁷ <i>E. coli</i>	SSH-G D VD YAQAS ELAKA IAQL RVIELTKKAM-----	

111 121 131

Fig. 7. Continued

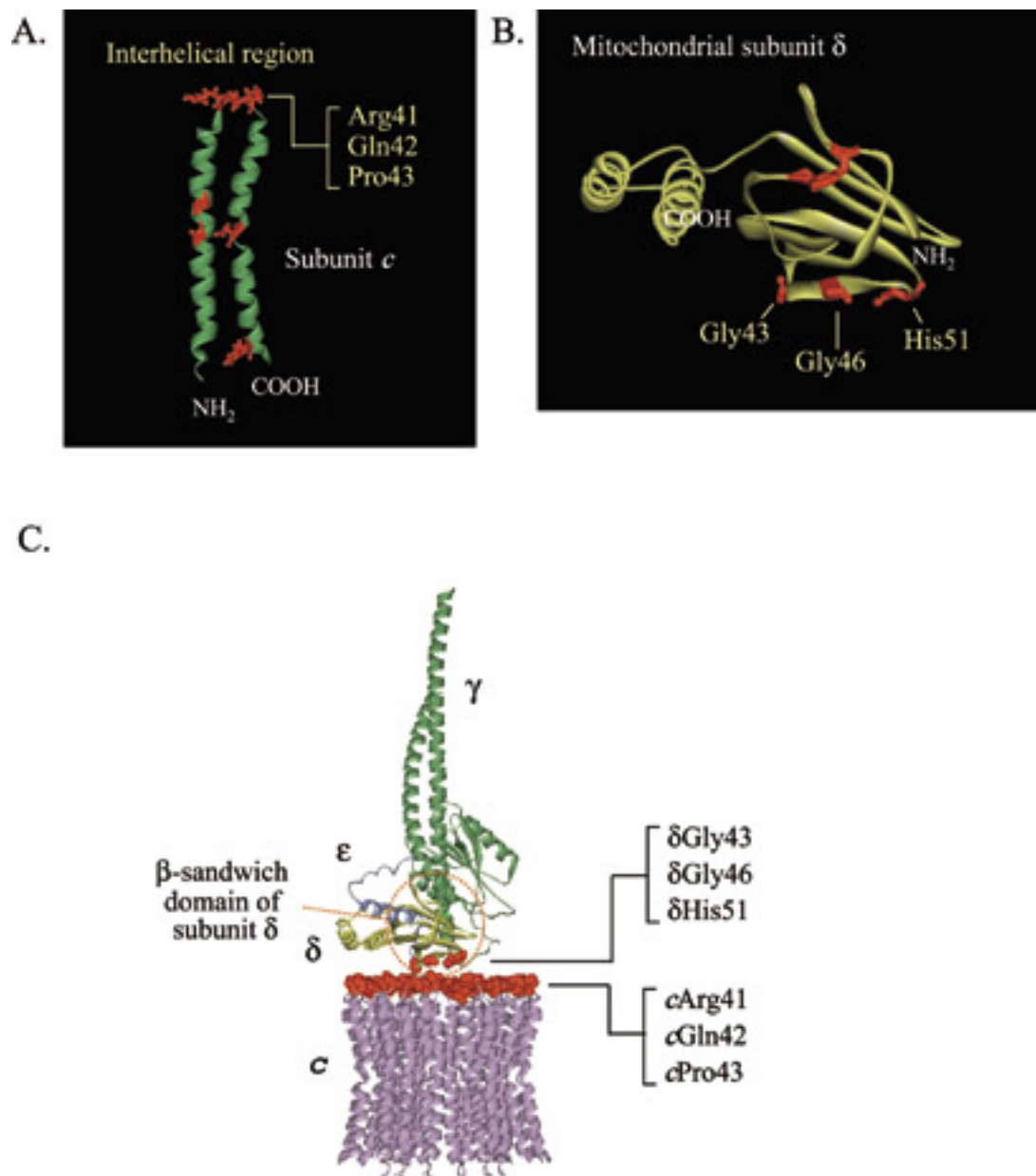


Fig. 8. Conserved interaction sites at the interface of rotors of F_1 and F_0 . The conserved residues in subunit δ (mitochondria) or ϵ (bacteria and chloroplast) and subunit c are numbered according to bovine sequence and *E. coli* sequence, respectively. (A) Location of the conserved residues in the structure of bacterial subunit c (1c17; Rastogi and Girvin, 1999). Completely and highly conserved residues are drawn in red. (B) Location of the conserved residues in the structure of mitochondrial δ (1e79; Gibbons *et al.*, 2000). The conserved residues colored in red are located mainly in the N-terminal β -sandwich domain. (C) Interaction of the conserved residues of mitochondrial δ (or ϵ in bacteria and chloroplasts) with the conserved cluster at the interhelix loop of subunit c .

presented). There is some sequence homology between the sequences from bacteria and chloroplasts, but the degree of conservation is fairly low. The sequence alignment of subunit b from nonphotosynthetic eubacteria shows that the homology is fairly weak. Only one residue, Arg36

(*E. coli* sequence), which is located at the region between the membrane spanning helical domain and the coiled-coil dimerization domain, is completely conserved. In contrast, the sequences of mitochondrial subunit b are fairly different from those from bacteria and chloroplasts. Apart from

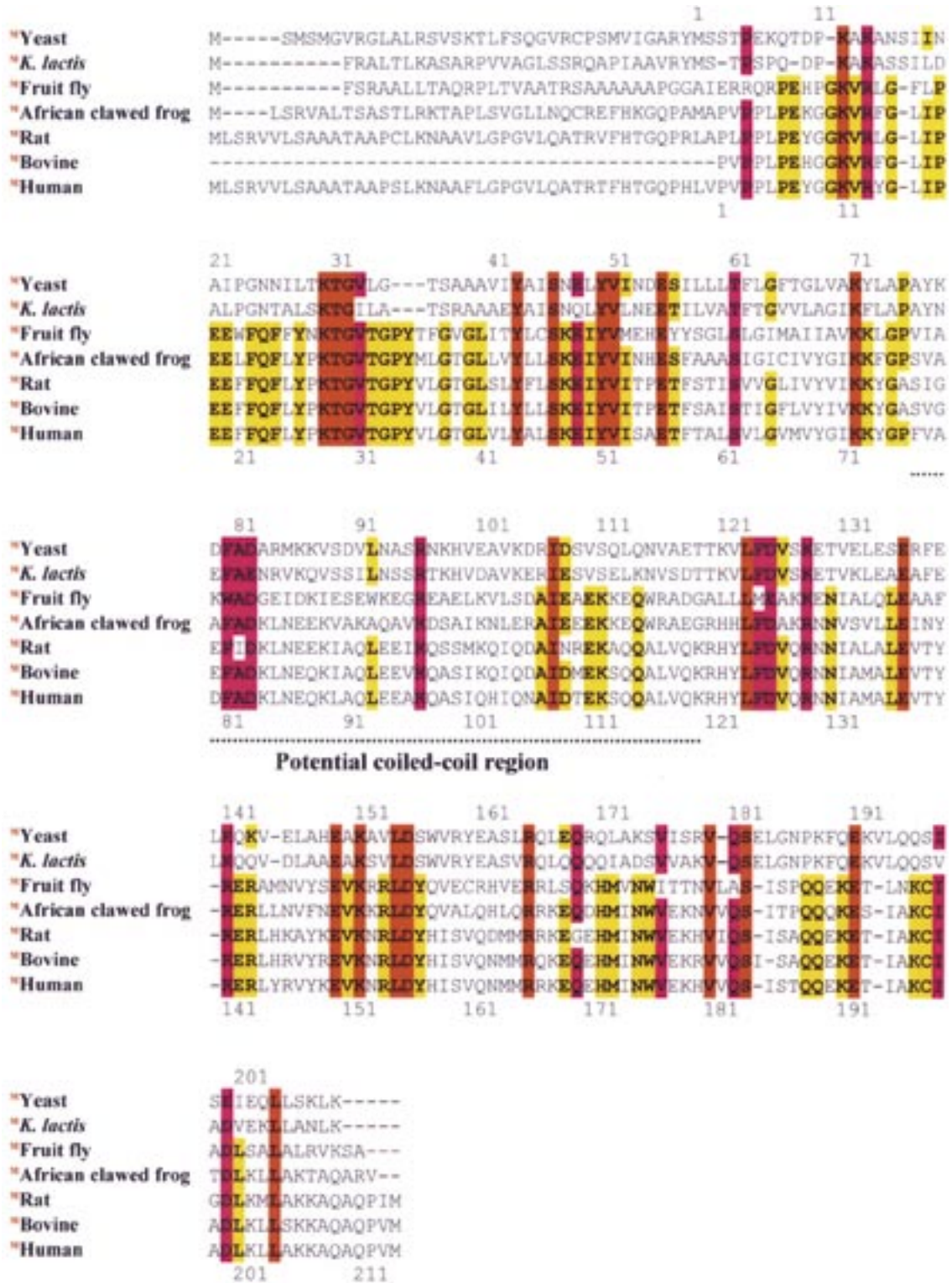


Fig. 9. Sequence alignment of mitochondrial subunit *b*. Conserved residues are highlighted the same as in Fig. 6. The potential coiled-coil region is marked by a dotted line.

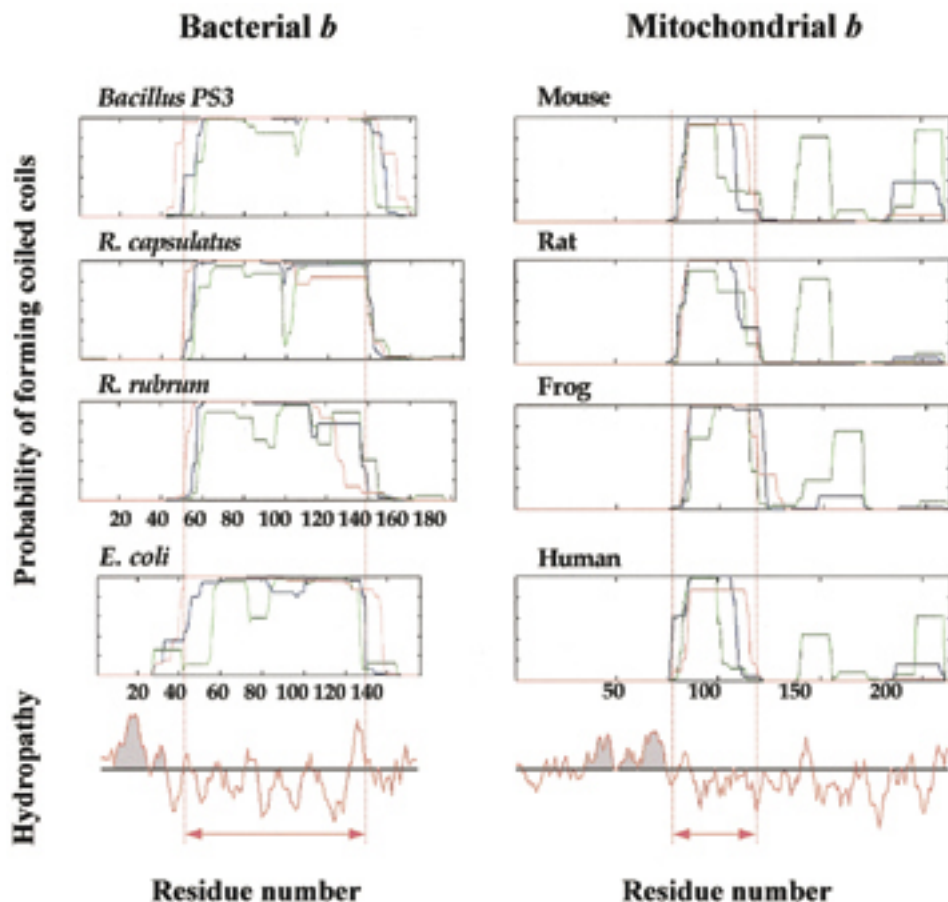


Fig. 10. Probability of forming coiled-coils and hydropathy plots for subunit *b* from bacteria and mitochondria. The dimerization domain of mitochondrial subunit *b* is predicted to be about half that of the bacterial *b* subunit. The predicted membrane-spanning domains are shaded in gray in the hydropathy plots. For the prediction of coiled-coil domain, the algorithm of Lupas *et al.* (1991) was used. It is based on the relative frequency of occurrence of amino acids at each position of the coiled-coil heptad repeat. For the hydropathy plot, the hydropathy index by Kyte and Doolittle (1982) was used.

the length of sequence, the sequence alignment of mitochondrial subunit *b* from seven different sources shows that more than 20 residues are completely conserved, and that the conserved residues are located fairly even along the sequence (Fig. 9).

In the sequence analysis of bacterial subunit *b* for the coiled-coil region, as shown in Fig. 10, the sequence from ~50 to ~130 (or 140) in the *E. coli* sequence is predicted to form a dimer, which is in line with the study on the dimerization domain of the *E. coli* subunit *b* mentioned above. In contrast, the coiled-coil domain for dimerization in the mitochondrial subunit *b* is predicted to be much shorter, although the length of the mitochondrial sequences is longer than that of bacterial sequences. Specifically, the dimerization domain of mitochondrial subunit *b* is predicted to encompass the sequence from about residue 80 to 120 in

the human sequence, which is about half that found for bacterial subunit *b*. From the sequence alignments and analyses of subunit *b*, it is interesting to speculate that the mitochondrial subunit *b* has maintained both high conservation of protein sequences and a shorter dimerization domain for the purpose of interacting with one or more supernumerary subunits.

NOTE

Just prior to the acceptance of this manuscript, a very nice experimental paper was published describing that the helices 1 and 5 in the N-terminal domain of bacterial δ , which is equivalent to mitochondrial OSCP, provide the binding surface of δ to F_1 [Weber, J., Wilke-Mounts, S. and Senior, A. E. (2003). Identification of the F_1 -binding

Surface on the δ -subunit of ATP Synthase. *J. Biol. Chem.* **278**, 13409–13416]. The conclusion in the paper obtained using the fluorescence spectroscopic analysis of the Trp residues in the δ subunit well supports our conclusion derived from the bioinformatic analysis of the sequences of mitochondrial OSCP and bacterial/chloroplast δ subunits, that the four-residue conserved cluster in the subunit plays an essential role in the interaction of the subunit with F_1 at the interface.

REFERENCES

- Altschul, S. F., and Konini, E. V. (1998). *Trends Biochem. Sci.* **23**, 444–447.
- Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J. (1997). *Nucleic Acids Res.* **25**, 3389–3402.
- Borghese, R., Turina, P., Lambertini, L., and Melandri, B. A. (1998). *Arch. Microbiol.* **170**, 385–388.
- Capaldi, R. A., and Aggeler, R. (2002). *Trends Biochem. Sci.* **27**, 154–160.
- Crofts, A. R. (1992). PSAAM—A Protein Sequence Analysis and Modeling Package Copyright, University of Illinois. Available from The Biotechnology Center, University of Illinois, 901 S. Mathews, Urbana, IL 61801.
- Dunn, S. D., Revington, M., Cipriano, D. J., and Shilton, B. H. (2000). *J. Bioenerg. Biomembr.* **32**, 347–355.
- Fillingame, R. H., and Divall, S. (1999). *Novartis Found. Symp.* **221**, 218–229.
- Gibbons, C., Montgomery, M. G., Leslie, A. G. W., and Walker, J. E. (2000). *Nat. Struct. Biol.* **7**, 1055–1061.
- Golden, T. R., and Pedersen, P. L. (1998). *Biochemistry* **37**, 13871–13881.
- Hirokawa, T., Boon-Chieng, S., and Mitaku, S. (1998). *Bioinformatics* **14**, 378–379.
- Hong, S., and Pedersen, P. L. (2002). *Arch. Biochem. Biophys.* **405**, 38–43.
- Joshi, S., Cao, G. J., Nath, C., and Shah, J. (1996). *Biochemistry* **35**, 12094–12103.
- Kramer, J. A. (2001). *Mol. Biotechnol.* **19**, 97–106.
- Krogh, A., Larsson, B., von Heijne, G., and Sonnhammer, E. L. (2001). *J. Mol. Biol.* **305**, 567–580.
- Kyte, J., and Doolittle, R. F. (1982). *J. Mol. Biol.* **157**, 105–132.
- Lupas, A., Van Dyke, M., and Stock, J. (1991). *Science* **252**, 1162–1164.
- Nakamoto, R. K., and Al-Shawi, M. K. (1995). *J. Biol. Chem.* **270**, 14042–14046.
- Noji, H., Yasuda, R., Yoshida, M., and Kinosita, K., Jr. (1997). *Nature* **386**, 299–302.
- Pedersen, P. L., Ko, Y. H., and Hong, S. (2000a). *J. Bioenerg. Biomembr.* **32**, 423–432.
- Pedersen, P. L., Ko, Y. H., and Hong, S. (2000b). *J. Bioenerg. Biomembr.* **32**, 325–332.
- Rastogi, V. K., and Girvin, M. E. (1999). *Nature* **402**, 263–268.
- Revington, M., Dunn, S. D., and Shaw, G. S. (2002). *Protein Sci.* **11**, 1227–1238.
- Revington, M., McLachlin, D. T., Shaw, G. S., and Dunn, S. D. (1999). *J. Biol. Chem.* **274**, 31094–31101.
- Richter, M. L., Hein, R., and Huchzermeyer, B. (2000). *Biochim. Biophys. Acta.* **1458**, 326–342.
- Schuler, G. D., Altschul, S. F., and Lipman, D. J. (1991). *Proteins* **9**, 180–190.
- Tanabe, M., Nishio, K., Iko, Y., Sambongi, Y., Iwamoto-Kihara, A., Wada, Y., and Futai, M. (2001). *J. Biol. Chem.* **276**, 15269–15274.
- Thompson, J. D. (1994). *Nucleic Acids Res.* **22**, 4673–4680.
- Velours, J., and Arselin, G. (2000). *J. Bioenerg. Biomembr.* **32**, 383–390.
- Wilkens, S., Dunn, S. D., Chandler, J., Dahlquist, F. W., and Capaldi, R. A. (1997). *Nat. Struct. Biol.* **4**, 198–201.
- Wishart, D. S., Stothard, P., and van Domselaar, G. H. (2000). *Methods Mol. Biol.* **132**, 93–113.